Reusable Influence Diagram Modules for Rapid Program Planning and Decision Analysis

Phase I SBIR Final Report
Contract: DAAH01-94-R296

April 20th, 1996

Max Henrion, Ph.D.,

Lumina Decision Systems, Inc.

Project performance: October 10, 1994 to April 9, 1995

Sponsor: Advanced Research Projects Agency
3701 N. Fairfax Drive
Arlington, VA 22203-1714

~~Limited to US Government Agencies Only~~

Lumina
DECISION SYSTEMS

4984 El Camino Real, Suite 105
Los Altos, CA 94022

henrion@Lumina.com

Tel. 415-254-0189 FAX 415-254-0292

Abstract

In an era of declining government budgets and increasing business competition, it is critical to base investment and planning decisions on careful analysis of the cost-effectiveness of available options. Building sound and comprehensible models to support effective decision making is currently a demanding and costly process. In this Phase I SBIR Project, we have developed tools that greatly reduce the difficulty and cost of applying these methods by developing an electronic library of reusable decision modules. Specifically, we have designed, prototyped, and demonstrated the practicality of a suite of tools to:

1. use influence diagrams modules as an intuitive graphical method for creating, storing, and communicating extensible, reusable decision models;
2. create an electronic library (Decision Resources On-line) of influence diagram modules and related information and software resources accessible via the World Wide Web,
3. retrieve, adapt, and assemble these influence diagram modules to create new decision models to address new decision situations
4. communicate the assumptions, and key sensitivities to provide qualitative insights into model recommendations

To demonstrate these tools, we applied them to develop a prototype resource to support decision making on a variety of important public health issues, including education, screening, testing, and treatment for breast cancer, Chlamydia, measles, and tuberculosis.

Anticipated benefits

The technologies developed during this project promise to make sound decision-analysis methods accessible, practical, and cost-effective for a far wider range of investment and planning decisions in US government and business than has been possible hitherto. Some of these techniques have been incorporated in Analytica™, commercial decision modeling software soon to be released by Lumina.

Key words

Economic models, decision analysis, influence diagrams, reusable software, World Wide Web, electronic libraries, public health.

Security classification: Unclassified

Availability: Unlimited

Number of pages: 40

Contents

If I have seen further than others, it is because I have stood on the shoulders of giants.

Sir Isaac Newton.

1. Project Summary and Objectives

In an era of declining budgets in government and increasing global competition for US business, it is critical to base investment and planning decisions on careful analysis of the costs and benefits of available options. Given the pace of technological change, it is essential to be able to address the uncertainties explicitly and to extend and reconfigure analyses rapidly and flexibly as the environment shifts and organizational objectives evolves. The goal of this project was to develop decision modeling tools to support these needs that are much dramatically easier, cheaper, and faster than current tools.

1.1 The cost of decision analysis and systems modeling

Decision analysis, coupled with systems modeling, provides a set of practical methods to support sound analysis for organizational decision making in complex and uncertain environments. These methods provide:

- integrated treatment of hard data from empirical measurements with soft data from expert judgments
- explicit analysis of risk and uncertainty about technical success, performance, schedules and costs using probabilities
- explicit modeling of multiple competing objectives, such as lifecycle costs, reliability, systems effectiveness and benefits
- clear visual display of qualitative structure of decision models using decision trees and influence diagrams. (See Figure 1.)
- effective communication to decision makers of the key assumptions, implications, and sensitivities to provide the qualitative insights that are the eventual goal of quantitative analysis.

These methods are increasingly widely used in government and business to address high-stakes decisions on a systematic basis (Henrion, Breese, & Horvitz, 1991). As yet, however, they have achieved only a tiny fraction of their potential in facilitating more effective and systematic decision making. A major obstacle to the wider use of these methods has been the scarcity and expense of trained decision analysts and modellers with the skills to create sound models from scratch for complex decision problems. Existing decision-analysis software can speed the process of model construction and analysis, but it still requires highly trained decision analysts for effective use. Moreover, in conventional practice, each new decision analysis is typically developed anew, rather than being built on previous work.

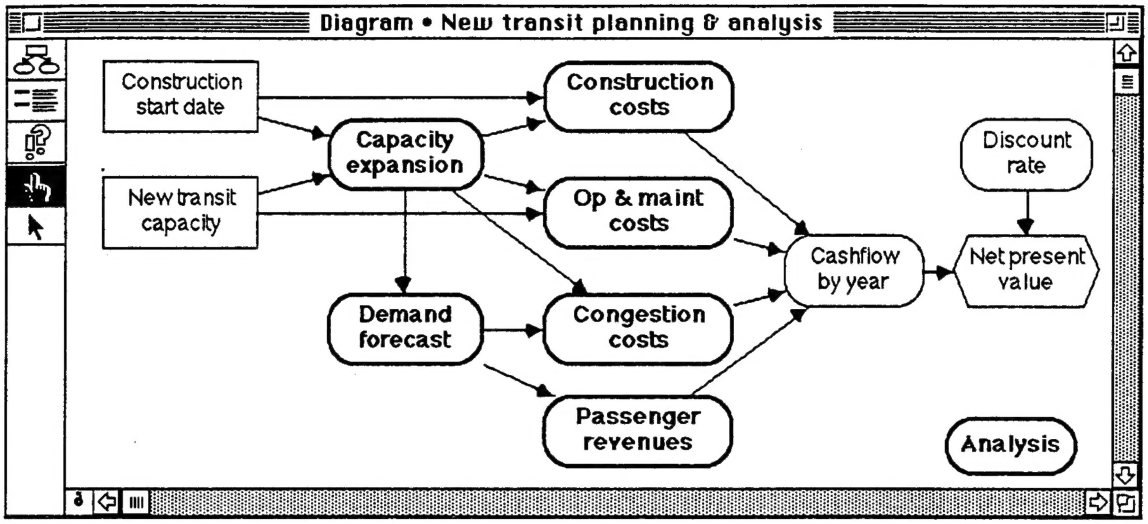


Figure 1a: An example influence diagram for decision analysis of options for the planned construction of a new rapid transit system. Influence diagrams provide an intuitive, graphical representation for creating and communicating the qualitative structure of decision models. See Section 3.1 for the interpretation of influence diagram notation.

1.2 Reusable influence diagram modules

Every organization faces classes of decisions that recur regularly for that organization — for example, a military transportation agency selects new airplanes and ground vehicles to acquire, an R&D organization selects a portfolio of R&D projects, or a public health agency selects measures to control a new disease epidemics. But, within a class of decisions, the specifics of each decision problem may vary in important ways from one to the next. A simple "canned" solution will not work well. Instead of having to create each decision model from scratch for each application, we have demonstrated how it is possible to start from a set of standard decision modules, which can be adapted and assembled to create new models. The objective of this project has been to develop and prototype methods to allow people to create their own decision-analysis models without requiring highly trained decision analysts, by reusing and adapting existing **influence-diagram modules (IDMs)**.

1.3 Decision Resource On-line Library

If modellers are to be able to build models from reusable modules, they must have easy access to a library of such components that address the kinds of problems of concern to them. The growing popularity and availability of the World Wide Web makes the Web the obvious candidate as the medium for accessing such a library. In this project, we developed a simple prototype of a library of influence diagram modules, available on the Web, which we termed **Decision Resource On-line Library (DROL).**

DROL also provides **adapters** , software agents to help the decision maker adapt each module to the specifics of the application, and **Assemblers**, software agents to combine multiple modules into a single decision model. These tools are

designed to allow analysts without extensive training in decision analysis to construct and apply decision-analysis models with much greater ease and speed than is possible with conventional tools.

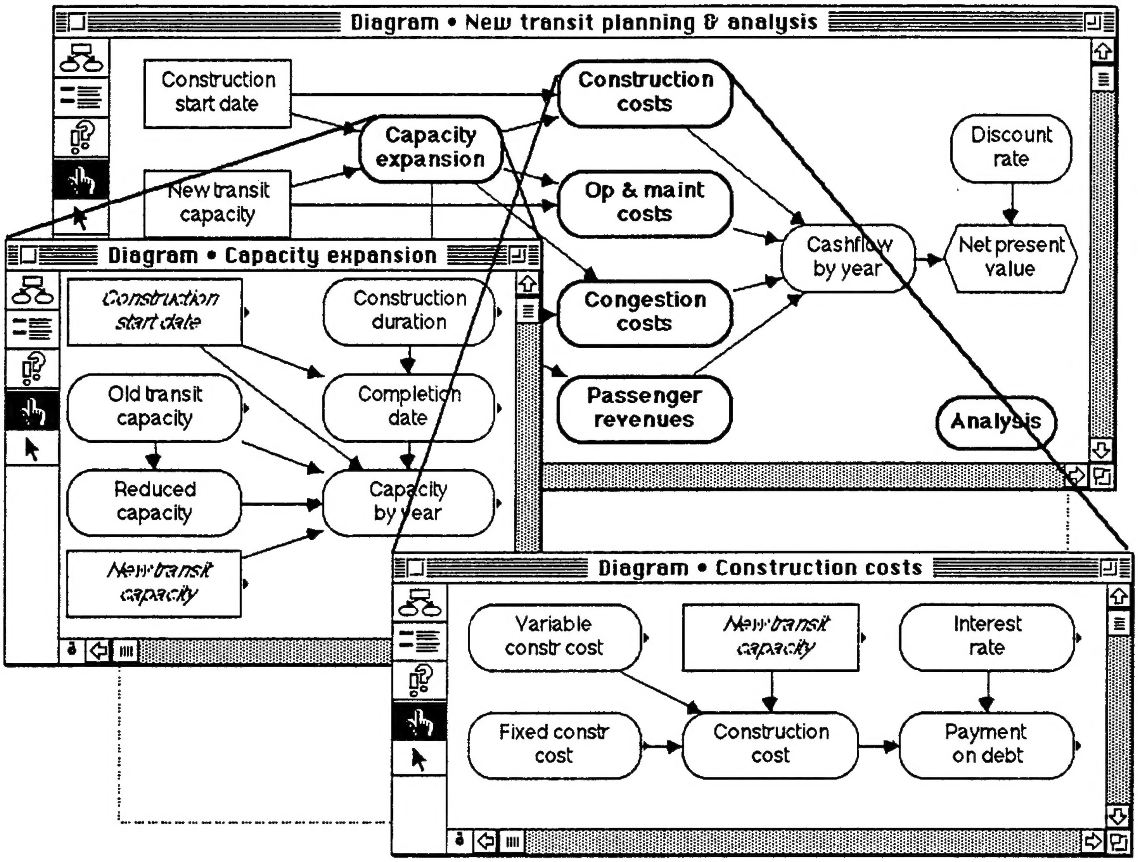


Figure 1b: Part of an influence diagram hierarchy showing the top diagram from Figure 1a, with details for two modules. Each node with a thick outline represents an influence diagrma module shown as a simple abstraction in the top diagram.

By providing a library of reusable IDMs, along with Adapters and Assemblers for such classes of decision, it should be possible for users with little or no training in decision analysis to create models tailored to a specific situation rapidly and at low cost. With such tools, it can become practical to apply such tools to a much wider array of situations.

1.4 Object-oriented decision modeling

Our notions of reusable influence-diagram modules (IDMs) borrow extensively from related concepts of reusable software components in software engineering, notably *object-oriented programming*, and *domain-specific system architectures* (Booch, 1991). Libraries of reusable software objects allow software engineers to build up new software much more rapidly and easily than they could by building software from scratch. Notable borrowings include the concepts of hierarchical module, interface specifications with public variables, object properties with inheritance, libraries of components, and graphical software tools

for examining and assembling modules. We term the approach we have developed **object-oriented decision modeling**.

1.5 HIPNet: Application to public health decisions

In our Phase I SBIR proposal, we planned to develop a specific application to an important class of decisions to demonstrate, and evaluate the potential of these techniques. At the suggestion of ARPA, we selected public health as the sample domain of application, and worked with a number of public health agencies to help identify important problems and realistic applications. Our protoype application is named **HIPNet (Health Information Planning Network).** HIPNet includes prototype decision modules for screening, testing, and treatment for breast cancer, Chlamydia, measles, and tuberculosis. The techniques and software we have prototyped are not specific to HIPNet and they could equally well be applied to facilitate decision modeling in a wide variety of other domains.

1.6 Summary of objectives

Our objectives during this Phase I project, were to design, prototype, and evaluate the practicality of a suite of tools to:

1. use influence diagrams as an intuitive graphical method for creating and communicating extensible, reusable decision models;
2. create an electronic library of influence diagram modules, related information and software resources accessible via the World Wide Web,
3. retrieve, adapt, and assemble these influence diagram modules to create new decision models to address new decision situations
4. communicate the assumptions, and key sensitivities to provide qualitative insights into model recommendations

1.7 Overview of report

In the next section, we introduce HIPNet as an illustrative application of the tools we have developed, including a scenario of how it might be used to improve decision making in response to the outbreak of a new disease, and example Web pages for the World Wide Web library responding to objective 2. In the following sections, we describe the tools we have designed to meet the remaining objectives listed above. Thus, Sections 3 through 6 describe the methods we have developed to extend the influence diagram notation to create IDMs, to adapt and assemble IDMs, and to communicate the assumptions, recommendations, and sensitivities of decision models. Each of these is illustrated by application to HIPNet. Next we discuss briefly the implementation of the prototype In the final section, we summarize the findings of our research, how well we have met our objectives, and directions for further research and development. We include a glossary of technical terms and acronyms at the end.

2. HIPNet: A sample application of DROL for public health decisions

The primary goal of this project was to design and evaluate innovative methods for building and applying decision models by adapting and assembling influence diagram modules obtained from an electronic library of modules. In order to evaluate these methods, we developed a prototype system to help decision making for issues in public health that we term **HIPNet (Health Investment Planning Network)**. HIPNet is intended as a pilot application to explore the possibilities and illustrate the potential value of DROL.

The intended users of HIPNet are public health decision makers at regional and local levels, such as directors of planning and policy for state and county departments of health, for public clinics, other health care providers such as health-maintenance organizations (HMOs), and, indeed, any organization concerned with maintaining and improving public health.

2.1 An illustrative scenario: Responding to the Siberian flu

In order to motivate the design of HIPNet, we developed a concrete scenario to illustrate one way in which HIPNet might be used. This scenario refers to an outbreak of a fictitious disease, the Siberian flu. We created it after brainstorming with public health experts. Our goal in providing the scenario was to help researchers, potential collaborators and users of such a system visualize how it might help a distributed network of organizations respond rapidly and effectively to a potential threat to society. The scenario is shown in Figure 2.

The scenario illustrates a number of important features of the approach. It emphasizes that the processes of identifying the problem, collecting information, conducting analyses, selecting appropriate responses, and implementing them are distributed geographically among a variety of organizations. It shows how analysts and modellers can make their expertise available to be applied in many different sites by providing decision modules and Adapter programs to apply the modules to specific regions or sites. It illustrates how these analysts can respond more rapidly by adapting an existing module for influenza rather than creating a new one form scratch. It underlines the importance of communication via electronic networks for sharing observations, decision models, and results. It also includes an evaluation phase in which analysts collect results from those who have made use of HIPNet to evaluate its effectiveness. It also emphasizes that HIPNet, like any new technology, will work best when properly integrated with existing technologies, such as EPI Info, and the host of public health resources available via the World Wide Web.

An illustrative scenario: Response of HIPNet to the Siberian Flu

- Health clinics in Oregon and California notice an apparent new strain of influenza and send reports to CDC via EPI INFO (CDC's existing network for exchange of epidemiological information)
- CDC investigators identify a hitherto unknown variant on the influenza virus, the Siberian flu, and work to assess the effectiveness of vaccines
- Several county public health officers send email requests to CDC asking for information to help them decide about public information and vaccination programs for the Siberian flu
- CDC takes an existing influence diagram module for flu, adapts it for the Siberian flu, using the best available estimates for transmission rates, morbidity, mortality, and the efficacy, costs, and risks of the vaccination.
- The HIPNet quality assurance team checks the module and Adapter for consistency with the best available data and estimates, and for usability, gives approval, releases to the HIPNet Library, and sends out an electronic announcement to concerned jurisdictions.
- State and county public health officers in the affected regions search the HIPNet library, download the module, and use the Adapter to tailor it to the local situation, including current prevalence, demographics, and public-health budget.
- Using the module, localities decide on public information campaigns, and vaccination programs for high-risk populations including seniors and school children.
- Local public health departments use EPI Info to forward information on their programs and on morbidity and mortality associated with Siberian flu to CDC.
- Using this information at the end of the epidemic, HIPNet analysts estimate that prompt action in affected regions has saved 8000 lives, 11,000 QALY (quality-adjusted live years), $230 million in healthcare costs, and $420 million in economic benefits due to avoided sickdays off work.

Figure 2: An illustrative scenario of how HIPNet might be used to respond to an outbreak of a fictitious disease, to develop, review, disseminate, and evaluate a new decision module

The scenario illustrates how the development of the library of decision modules will be partly *prospective* — that is, development of an initial set of modules to address obvious current decision issues within the domain — and partly *reactive* — that is, development or adaptation of modules in response to new problems, such as the threat of a new epidemic.

2.2 HIPNet Web pages

In the following pages, we present selected pages from the protoype HIPNet site on the World Wide Web. Figure 3 shows the front page of the Web site, providing access to a variety of information resources, both within HIPNet and links to external Web sites providing related information on public health. Figure 4 provides an introductory overview of HIPNet. Other pages and screens for HIPNet will be used throughout the rest of this report to illustrate the various facilities we have prototyped.

For this prototype, we focused on decisions on public education, screening, diagnostic testing, and treatment policies in four domains:

- Breast cancer: Public information programs for mammography screening
- Chlamydia trachomatis (CT) screening, testing, and treatment
- Measles vaccination programs
- Tuberculosis screening and treatment

For each of these areas, we developed prototype influence diagram modules based on available research literature. These decision models were based on existing decision analysis and cost-effectiveness models developed by CDC, universities, hospitals, and other research organizations, in some cases updated or generalized to be applicable to a wider range of sites.

Figure 5 lists the topics on which decision modules might be available. (Only those underlined were created in the prototype.) Figures 6 and 7 provide more background information on selected public health issues and specific decision modules.

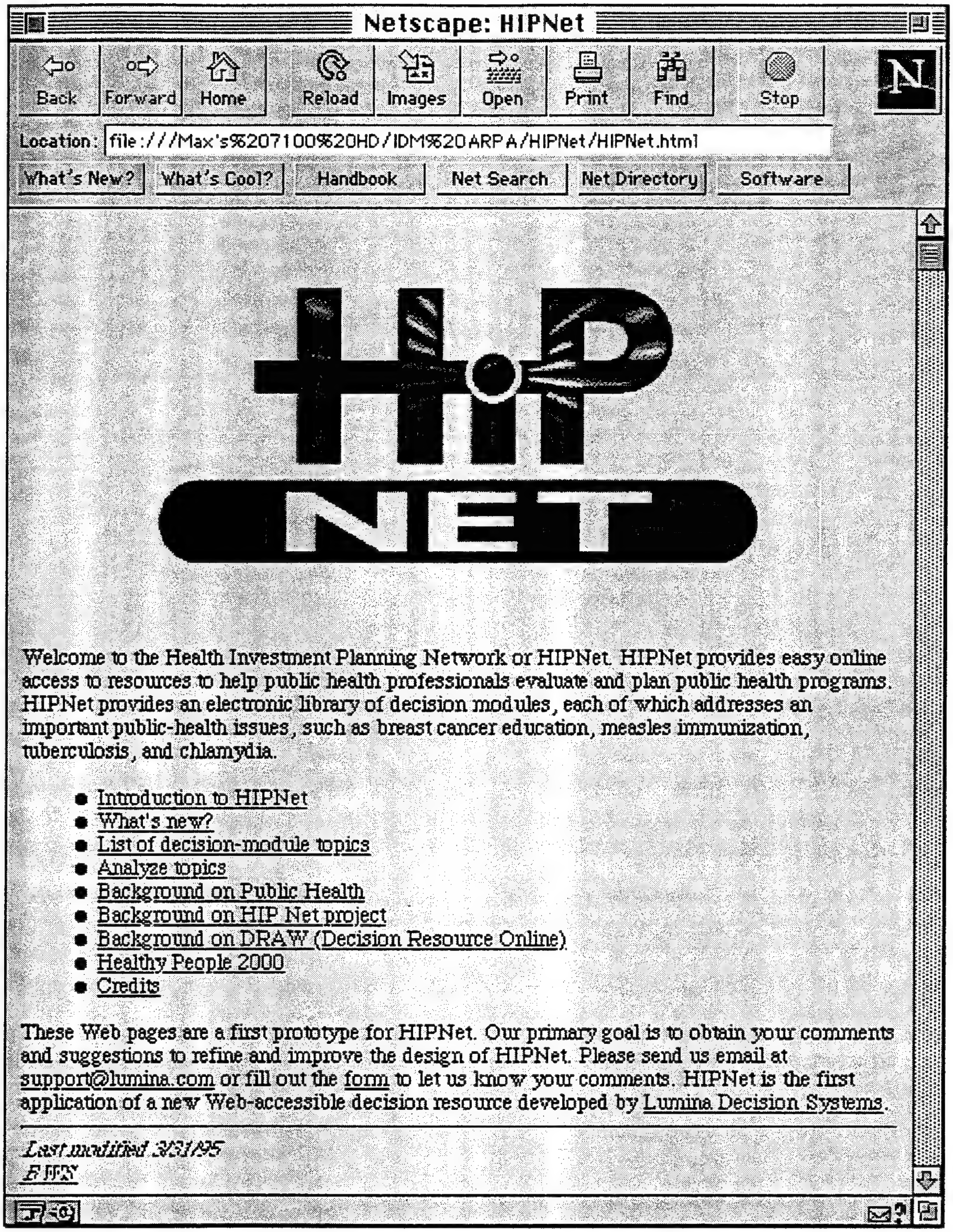


Figure 3: Front page of the prototype HIPNet Web site as seen through the Netscape browser for the World Wide Web.

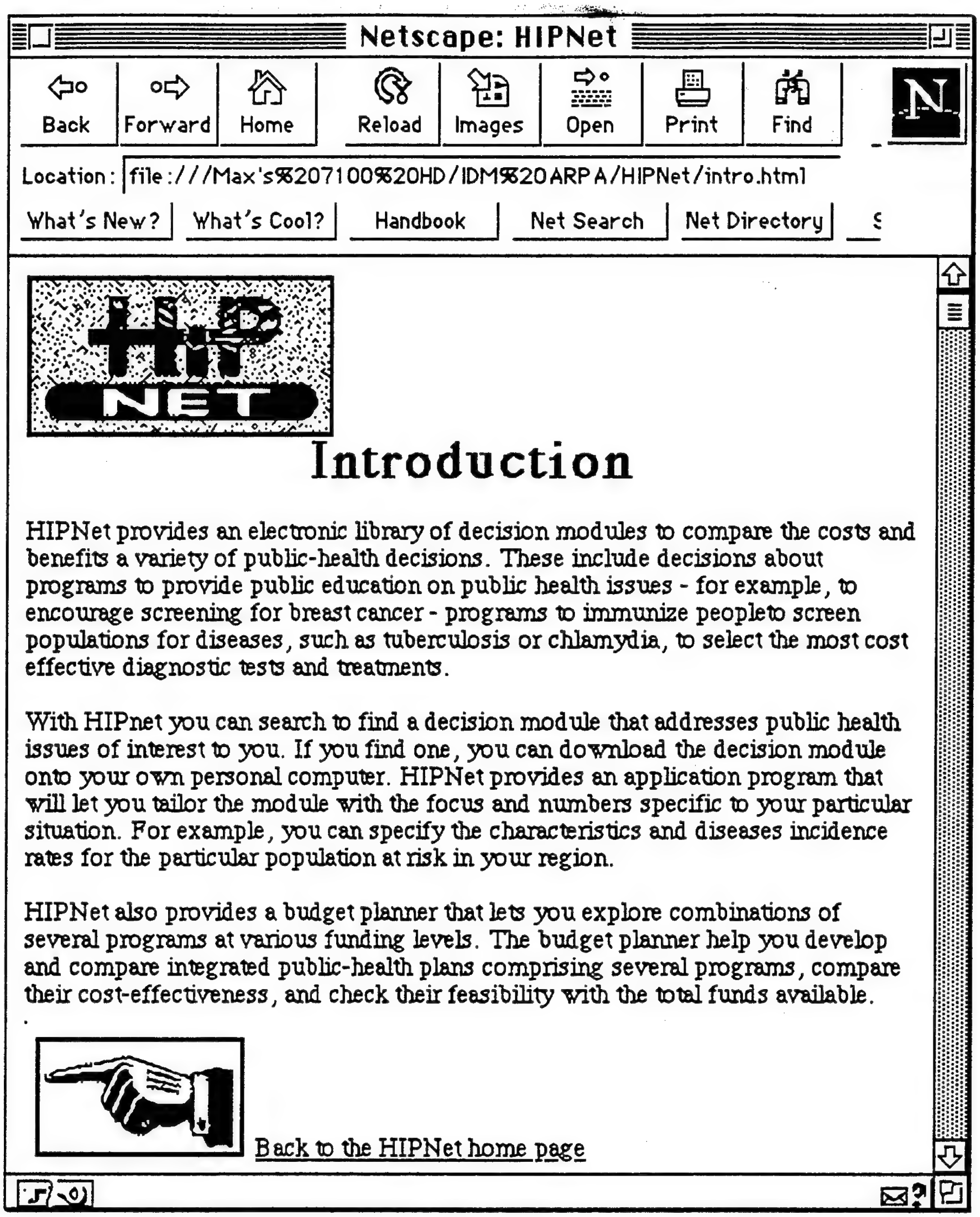


Figure 4: Introduction to HIPNet from prototype Web pages

Module Topics

A decision module is a model that lets you perform a decision analysis or cost-benefit analysis of one or more related public health issues. If you find a decision module for an issue of current concern to you, you can download it to your personal computer with a single click. HIPNet will help you apply and tailor it to the situation of concern to you. The following lists public-health topics. Decision modules are currently available for the topics marked by [??]. Click on one of those topics for details.

Preventive services

- Immunization and infectious diseases
- Cancer
- Sexually transmitted diseases
- Maternal and infant health
- Heart disease and stroke
- Diabetes and chronic disabling conditions
- HIV infection and AIDS
- Clinical preventive services

Health promotion

- Physical activity and fitness
- Nutrition
- Tobacco
- Alcohol and other drugs
- Family planning
- Mental health and mental disorders
- Violent and abusive behavior
- Education and community-based programs

Health protection

- Unintentional injuries
- Occupational safety and health
- Environmental health
- Food and drug safety
- Oral health

Figure 5: HIPNet Web page listing sample decision modules for public health applications. Only those items underlined exist in the prototype.

Tuberculosis Screening and INH Treatment regimen modules

Tuberculosis (TB) is a bacterial infection of the lungs. TB is a highly communicable disease spread through droplets breathed through the air. A typical infected person can infect five or more people if untreated. It is particularly dangerous to young children because of the danger of TB meningitis.

TB was one of the early success stories of public health. The combination of aggressive screening, followup and antibiotic treatments led to a dramatic decline in the number of cases (84,000 in 1953 to 22,000 in 1984). However, today vulnerable sub-populations in cities across the US are showing a resurgence in TB cases. The number of reported TB cases increased to 27,000 in 1992. The increase in TB is associated with many factors, including people whose immune system is compromised by AIDS, groups of immigrants from countries in which TB is endemic, and crowded living conditions in prisons, homeless shelters and the urban under class.

These modules focus on two aspects necessary to control TB. First, there is a module to determine an appropriate screening program for target populations. Second there is a module which evaluates the effectiveness of INH treatment regimens given variability in compliance and other factors.

Download the Tuberculosis screening module

Download the Tuberculosis treatment module

Figure 6: HIPNet Web page providing background information for the tuberculosis modules.

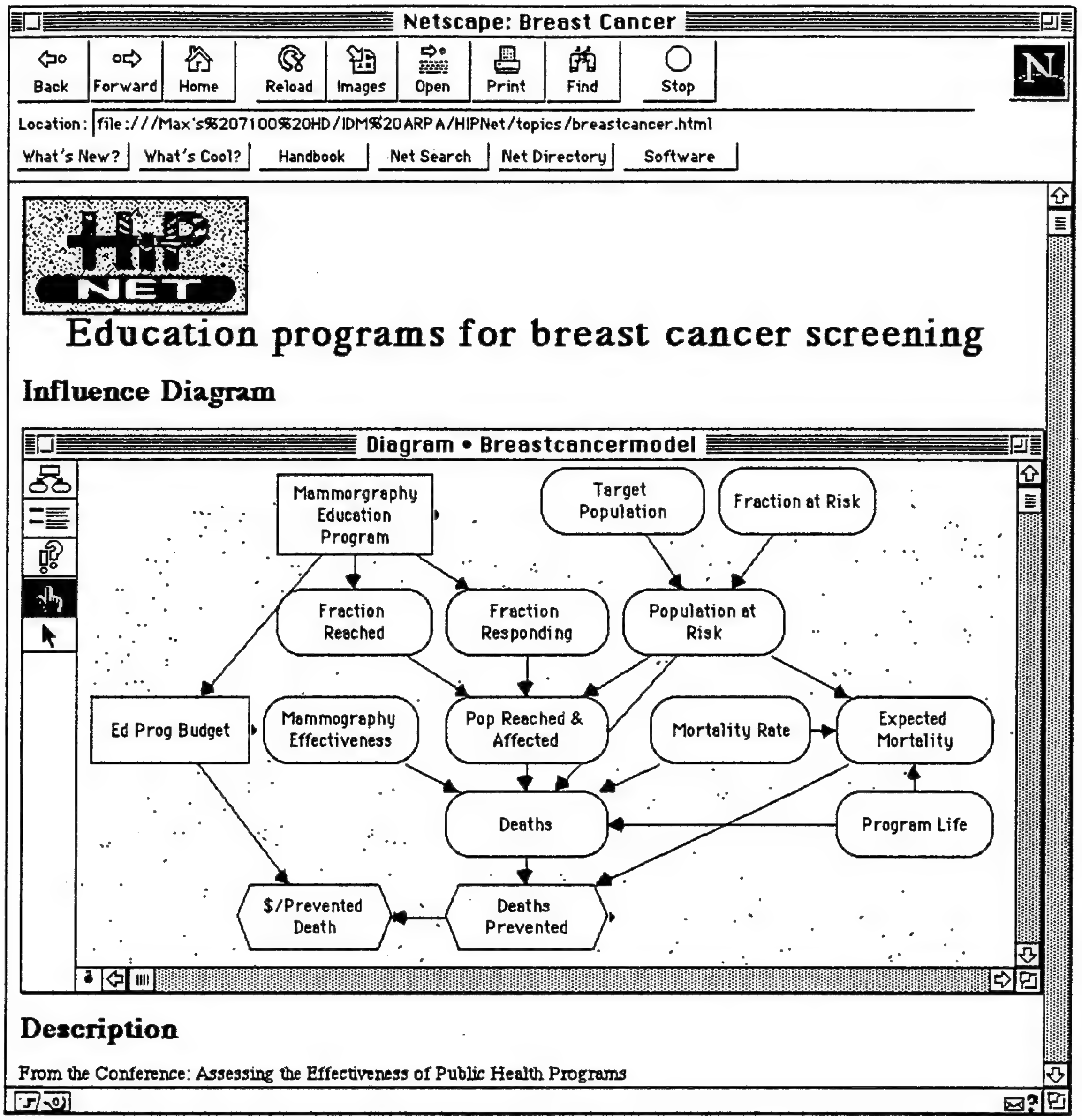


Figure 7: HIPNet Web page showing an influence diagram for the breast cancer screening module.

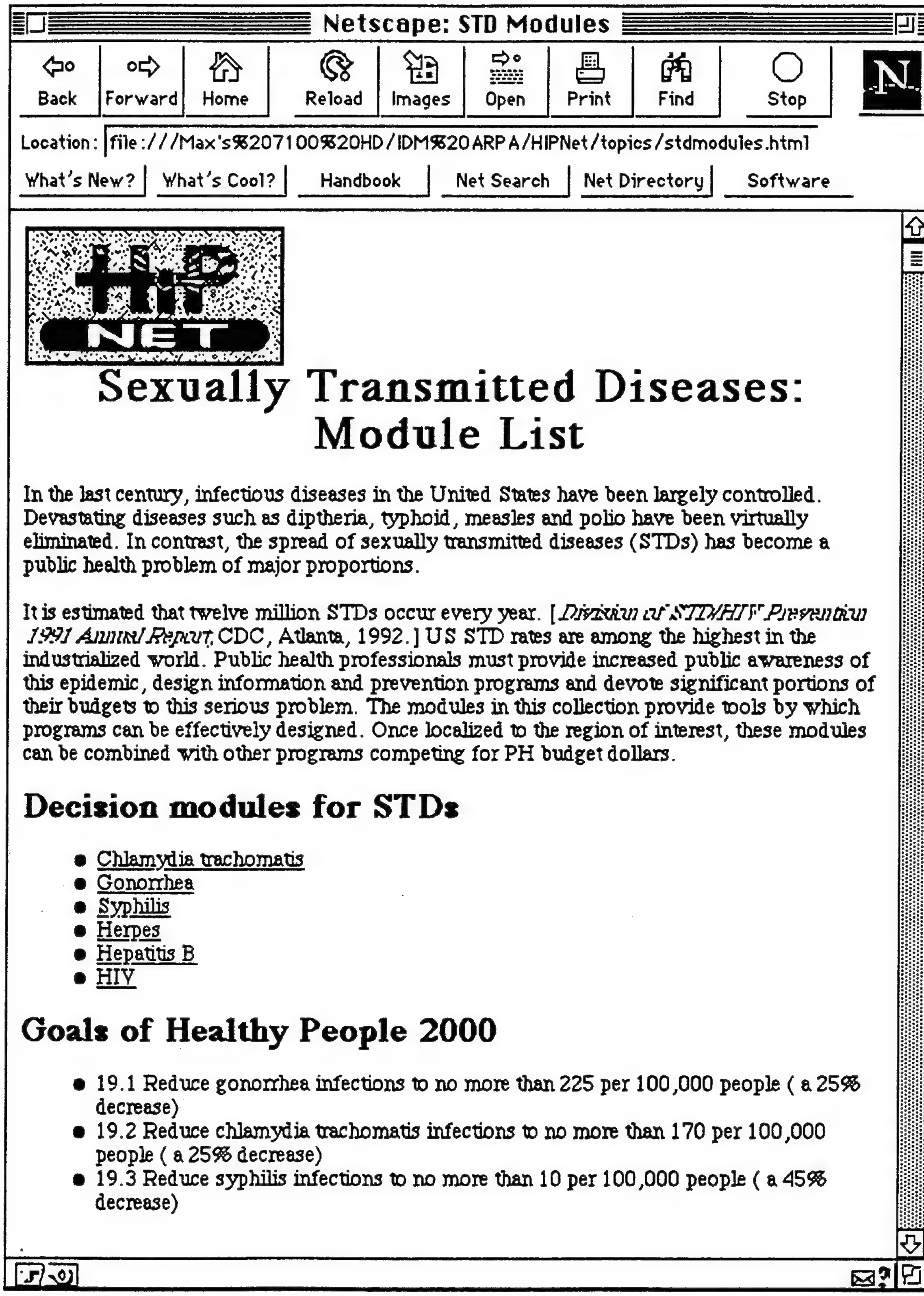
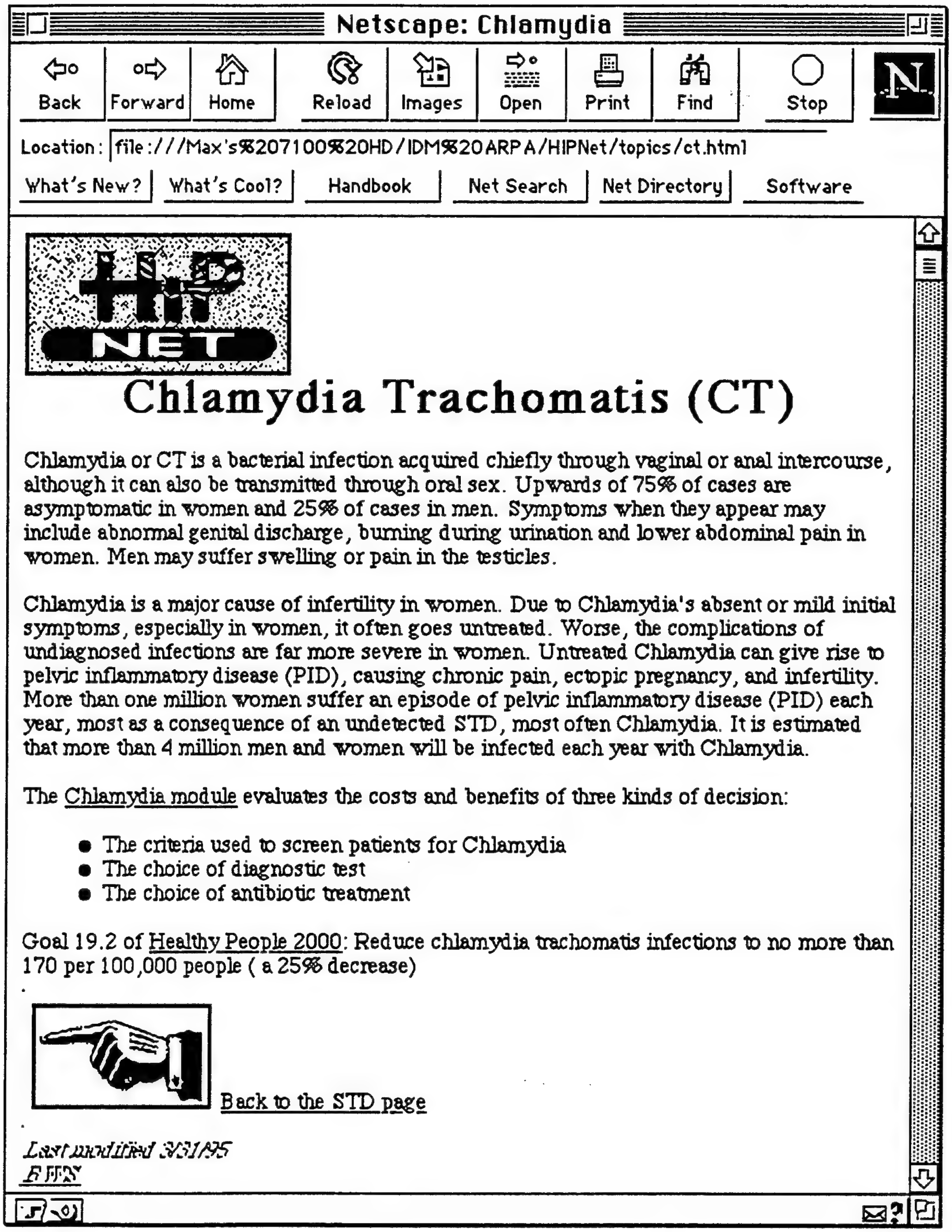


Figure 8: HIPNet Web page showing decision modules for screening and treatment of sexually transmitted diseases. Only the Chlamydia module exists in the prototype.


Netscape: Chlamydia

Back Forward Home Reload Images Open Print Find Stop

Location: file:///Max's%20700%20HD/IDM%20ARPA/HIPNet/topics/ct.html

What's New? What's Cool? Handbook Net Search Net Directory Software

Chlamydia Trachomatis (CT)

Chlamydia or CT is a bacterial infection acquired chiefly through vaginal or anal intercourse, although it can also be transmitted through oral sex. Upwards of 75% of cases are asymptomatic in women and 25% of cases in men. Symptoms when they appear may include abnormal genital discharge, burning during urination and lower abdominal pain in women. Men may suffer swelling or pain in the testicles.

Chlamydia is a major cause of infertility in women. Due to Chlamydia's absent or mild initial symptoms, especially in women, it often goes untreated. Worse, the complications of undiagnosed infections are far more severe in women. Untreated Chlamydia can give rise to pelvic inflammatory disease (PID), causing chronic pain, ectopic pregnancy, and infertility. More than one million women suffer an episode of pelvic inflammatory disease (PID) each year, most as a consequence of an undetected STD, most often Chlamydia. It is estimated that more than 4 million men and women will be infected each year with Chlamydia.

The Chlamydia module evaluates the costs and benefits of three kinds of decision:

- The criteria used to screen patients for Chlamydia
- The choice of diagnostic test
- The choice of antibiotic treatment

Goal 19.2 of Healthy People 2000: Reduce chlamydia trachomatis infections to no more than 170 per 100,000 people (a 25% decrease)

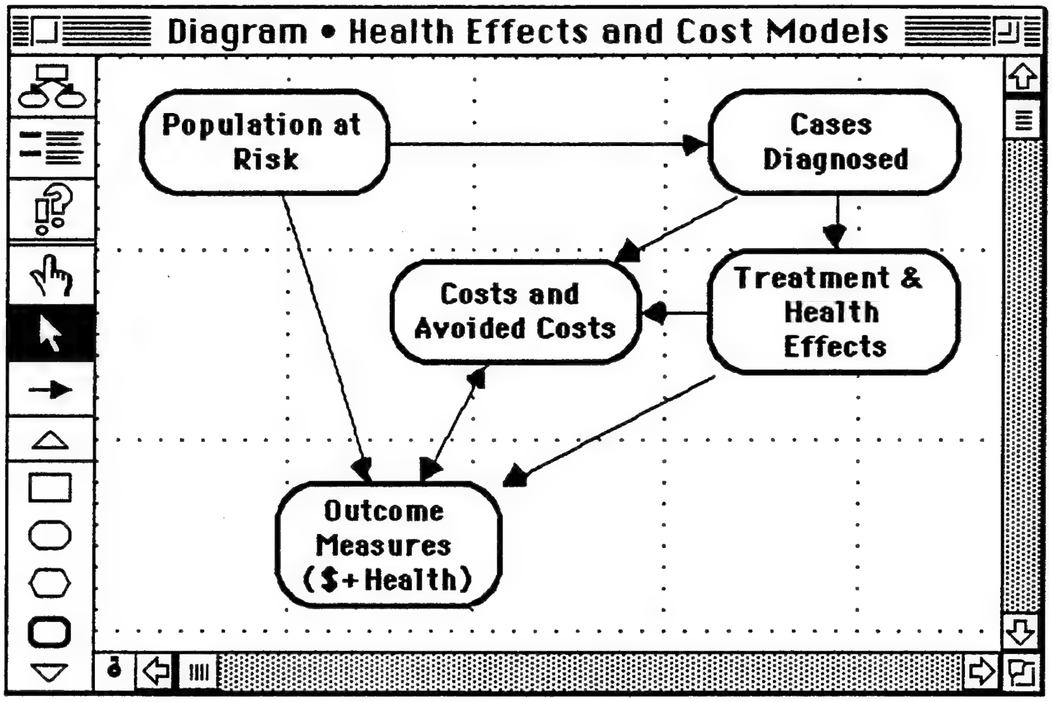
Back to the STD page

Last modified 3/31/95

Figure 9: HIPNet Web page providing basic information on Chlamydia trachomatis

CT Prevention Module

Influence Diagram



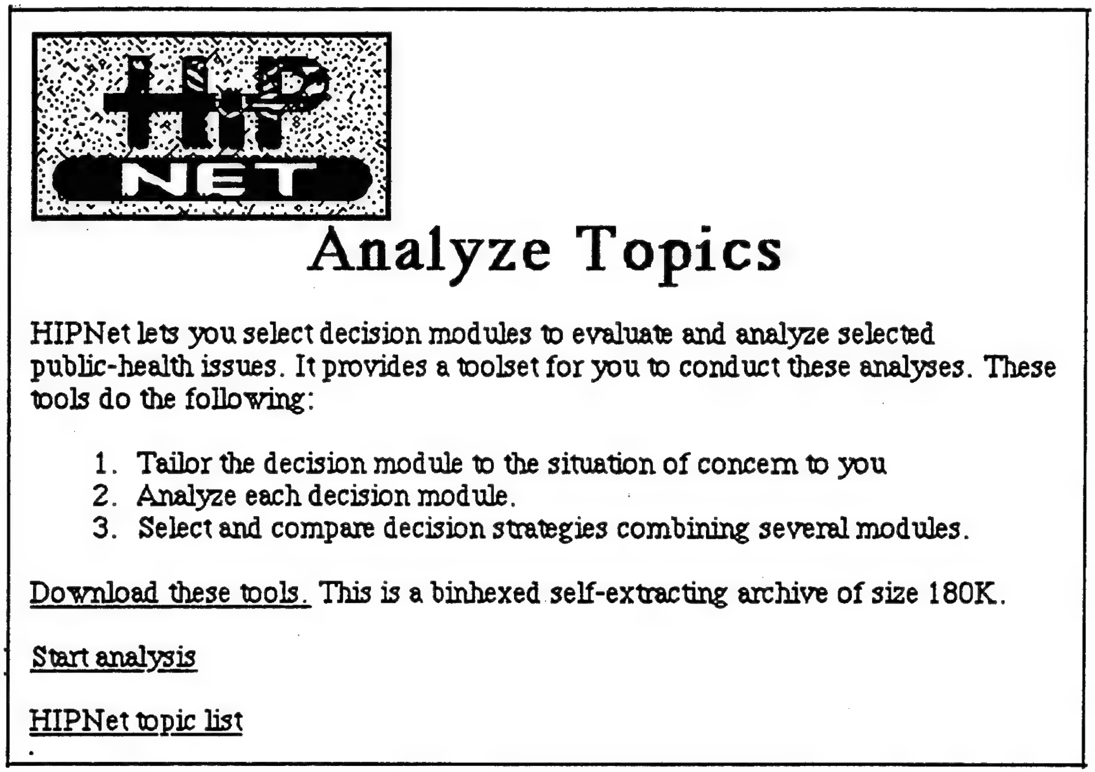
Description

The two major treatment regimines for Chlamydia trachomatis are Azithromycin (a single dose treatment) and a seven-day course of doxycycline. This module will take a look at the screening and treatment alternatives available to a public health decision maker.

Authors

Anne C. Haddix, Susan D. Hillis, and William J. Kassler
Transcribed by Bob Korsan and Max Henrion

Figure 10: HIPNet Web page providing top level influence diagram for screening, testing, and treatment of Chlamydia.



Analyze Topics

HIPNet lets you select decision modules to evaluate and analyze selected public-health issues. It provides a toolset for you to conduct these analyses. These tools do the following:

1. Tailor the decision module to the situation of concern to you
2. Analyze each decision module.
3. Select and compare decision strategies combining several modules.

Download these tools. This is a binhexed self-extracting archive of size 180K.

Start analysis

HIPNet topic list

Figure 11: HIPNet Web page explaining how to download selected modules and other resources for analysis.

By clicking on the download links, the user can cause HIPNet automatically to download the selected decision modules. Figure 11 describes how to download the related Adapter and Assembler information. The resources are automatically downloaded in compressed form (it takes less than 1 minute with a 14,400 baud modem line), expanded, and started. The downloaded programs are designed to appear like an extension to the Web pages for the user.

3. Reusable influence diagram modules (IDMs)

A major task in our Phase I research was to extend the conventional influence diagram notation to facilitate the creation and assembly of reusable modules based on influence diagrams. In this section, we first review the conventional influence diagram notation, and then describe the key extensions we have developed.

3.1 Conventional influence diagram notation

Influence diagrams were developed by decision analysts as a graphical representation for decision models. They have proved an intuitive but powerful notation for specifying and communicating the qualitative structure of decision models, providing a complementary representation to the more familiar decision tree. Influence diagrams show the essential variables and how they depend on one another.

- Rectangles denote the decision variables — those under control of the decision maker.
- Ovals denote chance variables — uncertain quantities not under control of the decision maker.
- Rounded rectangles represent intermediate deterministic quantities — which may depend on decisions and uncertain variables, but do not introduce additional uncertainty themselves. (An alternative convention is to use an oval with a double outline.)
- A hexagon denotes the objective or utility — a quantity evaluating possible outcomes of the decision and chance variables, and which the decision maker is trying to maximize. (Alternative graphic conventions sometimes used for objectives are the diamonds or rounded rectangle.)

Decision analysts prefer to quantify the objective in terms of utility, a numerical measure of the preference for an outcome, certain or uncertain, relative to other outcomes. Utility may be composed of multiple conflicting subobjectives or attributes — such as improved health, reduced mortality, reduced monetary costs, or reduced administrative complexity. Net benefits or net present value of benefits are also commonly used measures of objectives, particularly in societal and governmental decision making, and are sometimes viewed as societal utility functions.

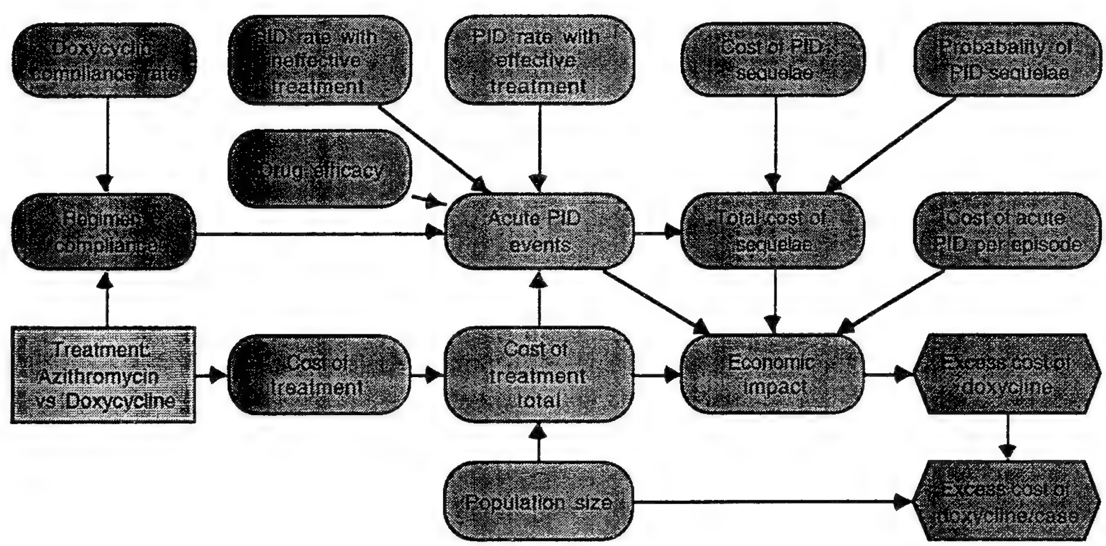


Figure 12: Influence diagram for the Chlamydia treatment module. This module compares the cost-effectiveness of Azithromycin and Doxycyclin, alternative antibiotics. Azithromycin is more expensive but requires only a single dose. Incomplete patient compliance with the full series of medication with Doxycyclin reduces its effectiveness.

3.2 Extensions of the influence-diagram notation

We have refined and extended the standard influence-diagram notation in a variety of ways, to facilitate the organization of models into hierarchical influence diagram modules (IDMs), that can be made available for reuse. We have incorporated these extensions into Lumina's decision-support software packages, Demos and its successor, Analytica. The following paragraphs describe these extensions:

Hierarchical modules: The rounded nodes with bold outlines denote IDMs included within the diagram — that is objects containing their own influence diagrams. Figure 13 reproduces the diagram from Figure 1b with two of its modules opened up to show their details. By means of this modular structure, one can organize a complex model containing hundreds or thousands of variables into a hierarchy of modules, each of which is small enough to be cognitively manageable and comprehensible.

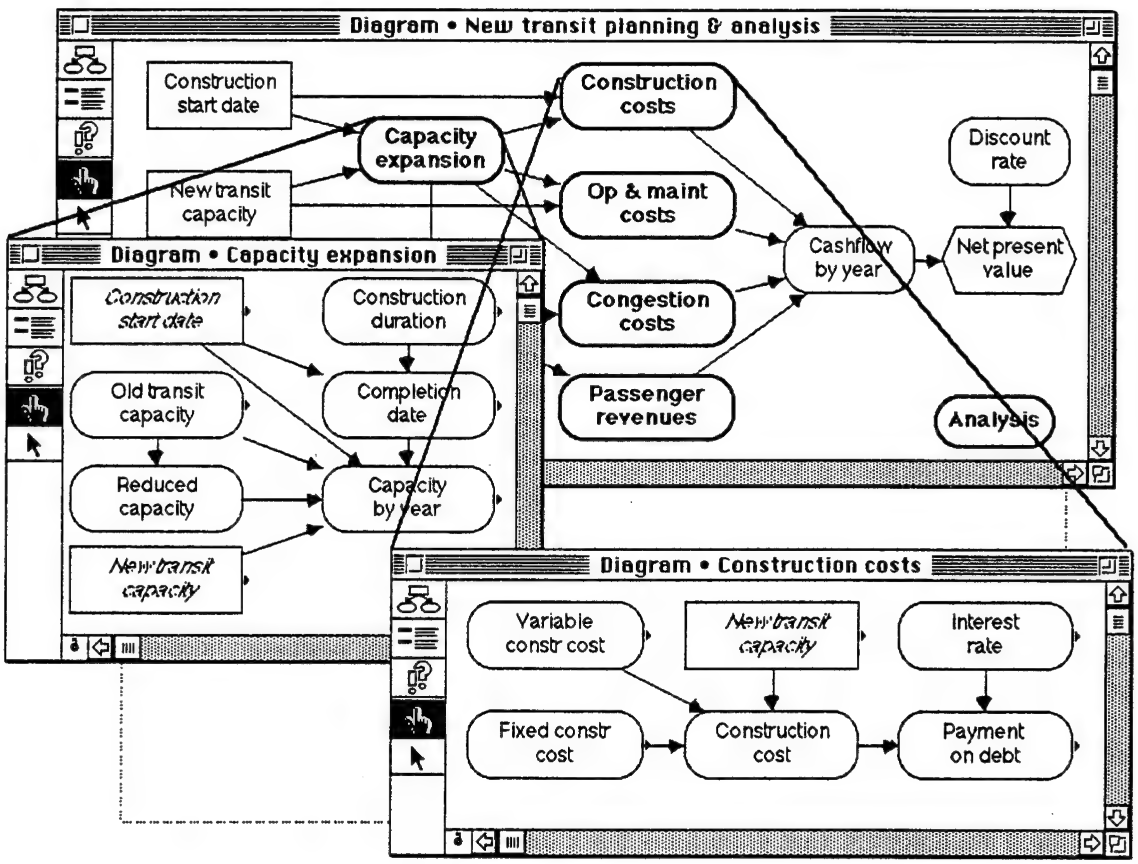


Figure 13: Part of an influence diagram hierarchy showing detailed for two modules of the diagram shown in Figure 1a. Each influence diagram module is an abstraction that encapsulates the details from the parent

Alias nodes show a variable that is actually contained in another IDM. Alias nodes are denoted by italic text. In Figure 13, the decision variables

"Construction start date" and "New transit capacity" in the top diagram, also appear as Alias nodes in the submodules shown below. Alias nodes provide a convenient way to show linkages between modules.

Public and private variables: The variables in each IDM can be categorized into inputs and outputs, that are public variables — available to other IDMs — and internal or private variables — not used by other IDMs. When linking IDMs, it is important that each input variable be provided a value by an output of another IDM. Some inputs may be provided as outputs from a user-interface module — or Form module — which allows the model user to supply the input value directly. Other inputs may be provided from other applications, such as databases.

Array values and Index variables: In Analytica, the value of any variable may be a simple scalar, certain or uncertain, or it may be an array with one or many dimensions. Each dimension of an array is identified by an Index variable. For example, in a healthcare model which handles different age-groups of patient — infants, children, teenagers, young adults, middle-aged, and elderly — these age-groups are identified by an index. For a dynamic model, the year of analysis — from 1995 to 2005, say — might also constitute an Index. Multiple decision options — such as, the alternative treatments, Doxycyclin and Azithromycin, for Chlamydia — also constitute an index. Index variables are shown by a parallelogram.

Required index variables: Some input dimensions may be required by inputs. For example, the Chlamydia modules is designed to compare the cost-effectiveness of two treatments, Doxycyclin and Azithromycin. It expects the treatment cost and effectiveness to be dimensioned by this Treatment index. In this way, the Treatment index forms part of the input specification. However, the module is flexible, in that it would be possible to extend the Treatment index, including additional treatment options. If the Treatment index is so expanded, all inputs dimensioned by Treatment will expand correspondingly. The user will need to provide appropriate values (costs and effectiveness) for the corresponding input tables. Thus, the Treatment index is required for these inputs, but the number and identity of values for the index can be modified.

Optional index variables: In many computer languages, it is a major headache to handle and match arrays with different dimensions. Typically, we want an IDM to generalize easily to handle inputs with new dimensions. For example, if a module can analyze options for a single health clinic, it should generalize easily to handle options for a group of clinics — given that the relevant inputs, such as patient numbers, disease prevalence rates, that vary from one clinic to another can be provided as arrays of numbers indexed by clinic. We have developed Analytica so that it automatically generalizes in this way. All arithmetic and other operations in the modeling language that work on scalars will automatically generalize to work on single or multidimensional arrays. In this way, there is no need to specify the dimensions for inputs that are not specifically

required by the module. Such dimensions will automatically propagate to generate outputs with similar dimensions.

Libraries: A library is a special kind of module that contains a set of user-defined functions and related variables. The main difference from other modules is that the list of functions and variables gets added to the Definition menu of Analytica, and so provides a set of extensions to the built-in functions and system variables. By use of Libraries, we can develop versions of Analytica with functions tailored to specific classes of application.

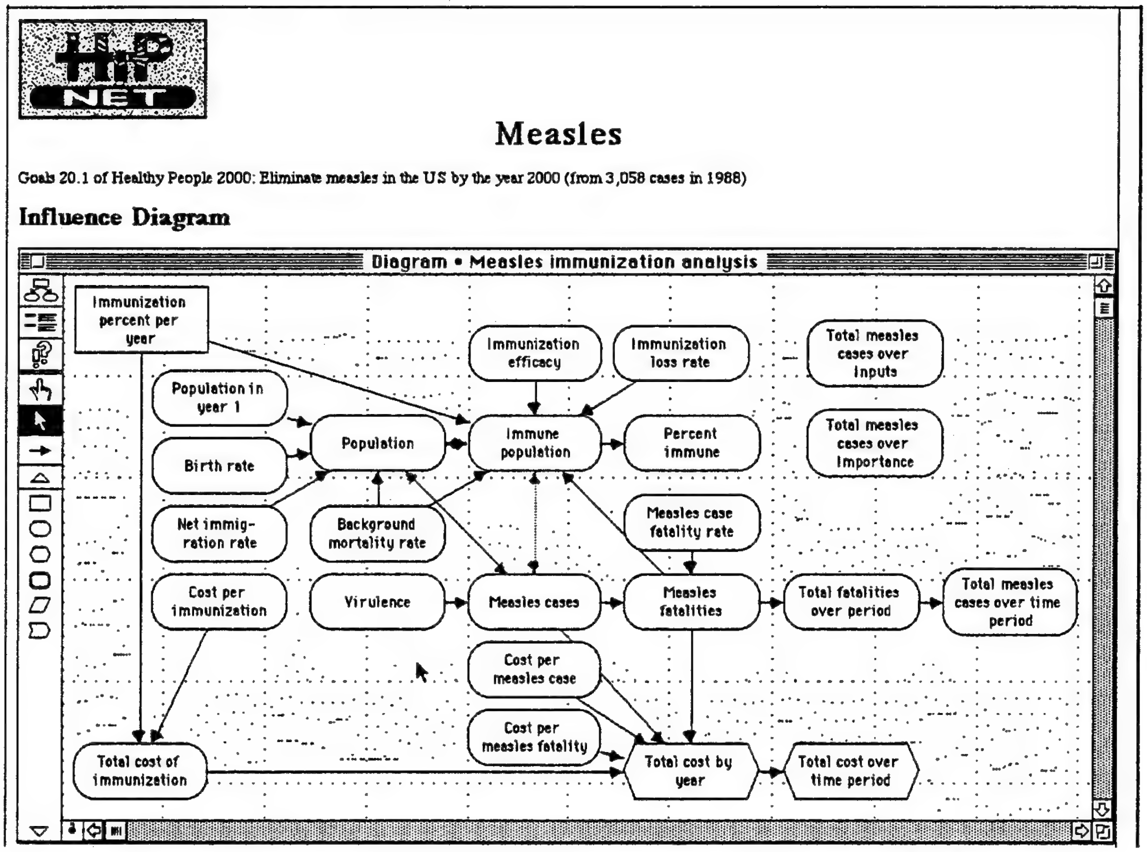


Figure 14: HIPNet Web page showing an influence diagram for the Measles module.

3.3 Linking and combining modules

The extensions to influence diagrams described above facilitate the creation of IDMs that can be stored separately and combined for new applications. The clear identification of inputs and outputs helps in defining these links. Each input must be provided with a value from an output of another module, or it must be provided a value directly by the model user. Outputs may, but need not, be linked to other variables.

Linking: We have implemented a variety of methods for combining modules in Analytica. The Link option in the File menu, lets the user add in a module to the

current model. The module may be added in to the current model, and become part of the model file, or it can be referred to by a simple link to the existing module file. The former method allows the user to modify the module within the current model without affecting the original module. The latter method allows the user to propagate changes back to the original module.

Merging: The module may be merged so that it automatically overwrites and updates any existing module or variables with the same name as the new module and its variables. By this means, it can modify those variables for its needs. By default, the software will flag name conflicts and ask the user how to resolve them.

Undefined variables: Analytica maintains a list of variables that have not been defined yet (or whose definitions contain missing elements), and displays it to the user. In this way, the modeller can keep track of input variables which still require values to be supplied either by other modules still to be added or directly by the user.

Propagation of array dimensions: The "intelligent arrays" design checks that required index variables are provided where necessary. It allows additional index variables specifying additional dimensions to be provided to inputs without requiring any modification of the downstream modules. All downstream variables will automatically be generalized to include this extra dimension in their results, displayed in table or graph form. This provides significantly greater flexibility than standard object-oriented languages, such as C++, and others, which cannot easily handle arrays whose number and identity of dimensions may change at link time or run time. For example, a module that performs a cost-benefit analysis of a drug for a single patient, will automatically generalize to perform the analysis for a set of drugs for a set of patients, provided the drug-specific and patient-specific input parameters are provided for each drug and patient, respectively. All those variables downstream will automatically be generalized to be indexed by the drug and patient Index variables.

3.4 Additional design features

The design features described above to support modular influence diagrams have all been implemented in Analytica, supported in part by the Phase I SBIR. We have identified a number of features as desirable to further support modularity and reusability, but as beyond the scope of our Phase I project.

Name scoping: Analytica does not currently support truly private variables. Variables can be accessed by name from any other module. Since this feature is standard in most computer languages, using name scoping, we did not regard it as worthy of research in this project. However, we regard it as an important addition for a future version.

Units checking: Currently, each variable has a field specifying its units of measurement. An important method for checking the validity of a model, and of

linkages between variables in different modules is compatibility of units. For example, a variable with units "Miles/gallon" cannot be validly linked with a variable in "Gallons/mile". On the other hand, variables with similar dimensions but different units — such as, "Miles/gallon" and "Kilometers/liter" — can be linked with provision of an automatic conversion factor. The check for compatibility of dimensions provides an extremely powerful semantic check on validity of linkages. The automatic conversion of compatible units is a convenience that increases the reusability of modules. The implementation of these features was beyond the scope of the Phase I project.

4. Tools to adapt and assemble modules

DROL provides basic library facilities to store, search, and retrieve IDMs via the World Wide Web. In order to make effective use of these IDMs, they must be adapted and linked together to fit the needs of a particular situation. A skilled analyst might be able to adapt and link modules working directly on the IDMs, but this would require considerable expertise, as well as understanding the details of the modules. To support use by nonexperts, we have developed prototype Adapter and Assembler software that guides the user through the process, requesting specific inputs values needed, and making linkages automatically between modules.

For example, the Chlamydia prototype module was developed based on work by researchers at CDC and the University of Washington, using numbers for disease prevalence rates, performance of diagnostic tests, and the cost and effectiveness of treatments specific to the populations and clinics studied. To apply this module for another site — for example, for a family planning clinic in downtown Los Angeles — requires modification of these input numbers. Indeed, since the patient information obtained for screening patients is different in Los Angeles than in Seattle, Washington, it is also necessary to modify the structure of the screening component.

Figure 15 shows the key menu options for the Adapter and Assembler agents, allowing selection of modules, supplying local conditions, reviewing assumptions, evaluating costs and benefits for individual modules, and assembling multiple decision modules for creating and comparing multiprogram strategies.

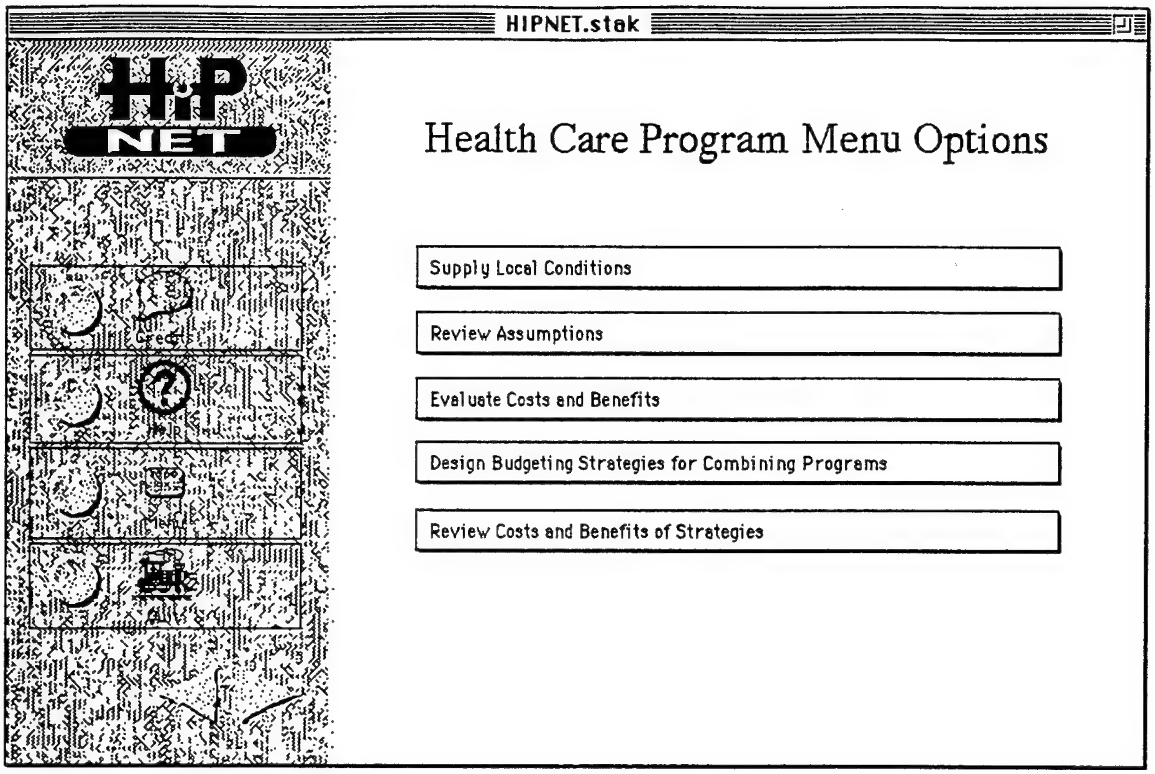


Figure 15: HIPNet screen listing menu options for adapting, reviewing, evaluating, and assembling decision modules.

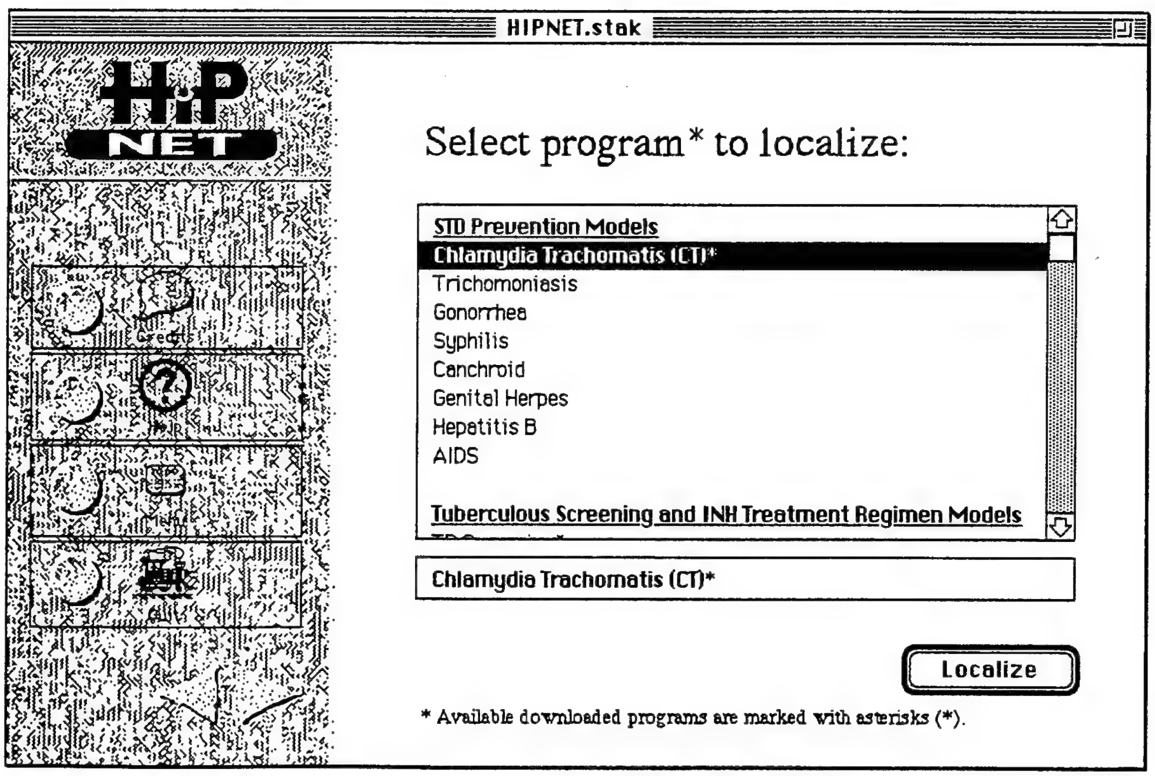


Figure 16: HIPNet screen to select decision modules for adaptation and assembly.

4.1 The Adapter program

HIPnet provides prototype *Adapters* for the Chlamydia and Breast Cancer modules which helps the user configure those modules for the specific location and situation. Thus, the user need know nothing special about the module, but merely answer the questions posed by the Adapter. The Adapter is a relatively simple program that requests these inputs and provides them to the module as an input file. Figures 17 to 19 shows three screens of an Adapter for Chlamydia.

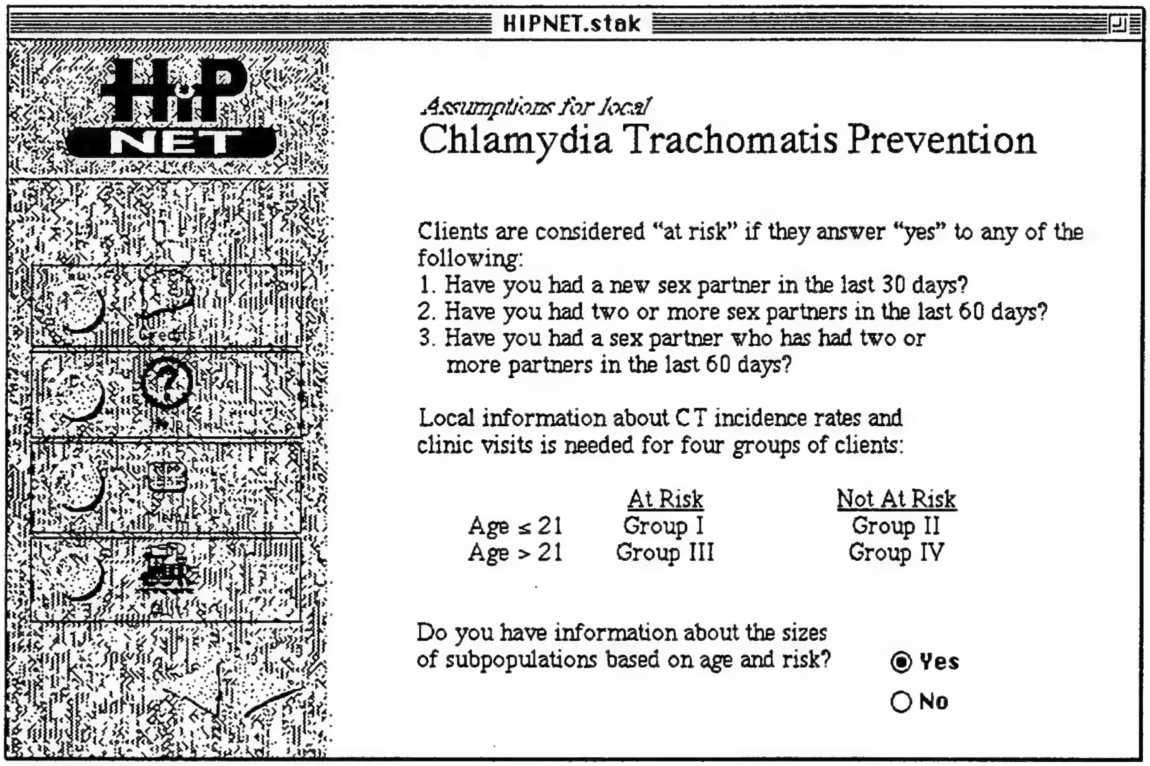


Figure 17: HIPNet Adapter screen for Chlamydia module to define at-risk groups for screening.

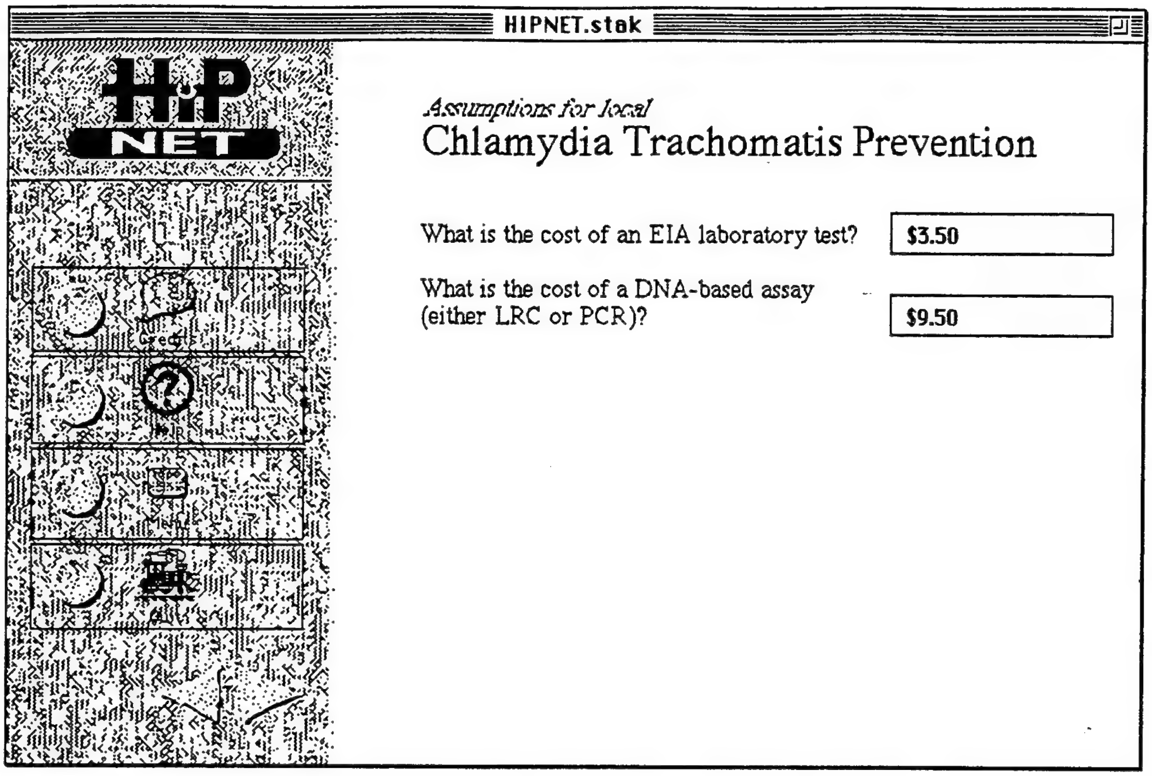


Figure 18: HIPNet Adapter screen for Chlamydia module to define costs of diagnostic tests.

4.2 Assessing uncertainty

Many of the quantities needed for decision making in public health, as in most other areas, are uncertain. For example, the prevalence rate of a disease in a population is not known exactly, but estimated from limited data. The data may be out-of-date, nonrepresentative, and not reliably coded, or inaccurate for a host of other reasons. Or there may be no numerical data, so the user is forced to provide a judgmental estimate. In any of these cases, the user can provide a range of values to be interpreted as a probability distribution to express the uncertainty about each quantity.

The prototype Adapter that we developed allows the user to express uncertainty explicitly about those input values that often have large uncertainties with significant importance to the results, such as the diseases prevalence rates. Figure 19 shows a screen that requests, low, high, and best estimates of a quantity. HIPNet interprets these as the minimum, maximum, and mode respectively of a triangular probability distribution. We have designed the Adapter to ask these three quantities in this sequence, with the best estimate last, because experimental research has shown that initial focusing on the central estimate, can bias the low and high values to be too close, resulting in an overconfident distribution.

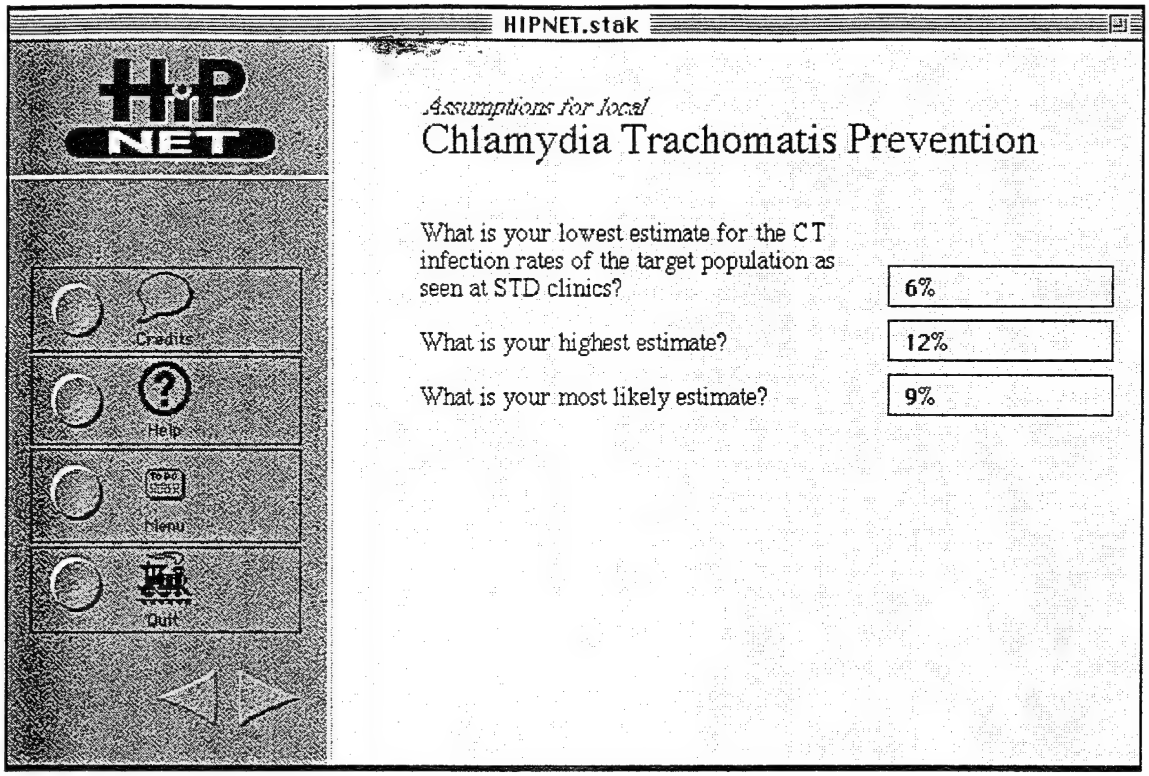


Figure 19: A screen from the HIPNet Adapter for the Chlamydia module asking the user for an uncertain input quantity to localize the module for a specific site. It requests low, high, and most likely estimates to express uncertainty about the infection rate in the target population, interpreted as the minimum, maximum, and mode of a triangular probability distribution.

The modules, using Demos, propagate these probabilistic values through a model to generate probability distributions for the outputs of interest, such as net benefit or cost-effectiveness.

4.3 The Assembler for a strategy table

An Assembler is a program, or agent, that helps the user combine and link selected IDMs to address a particular problem that requires more than one IDM. The Assembler identifies the required IDMs, creates links among them, and creates output views to display and analyze the results. The Assembler prototype that we have developed helps the user create a strategy table to analyze alternative strategies combining sets of public health options for several different public health programs. Figure 20 shows an example strategy table. A

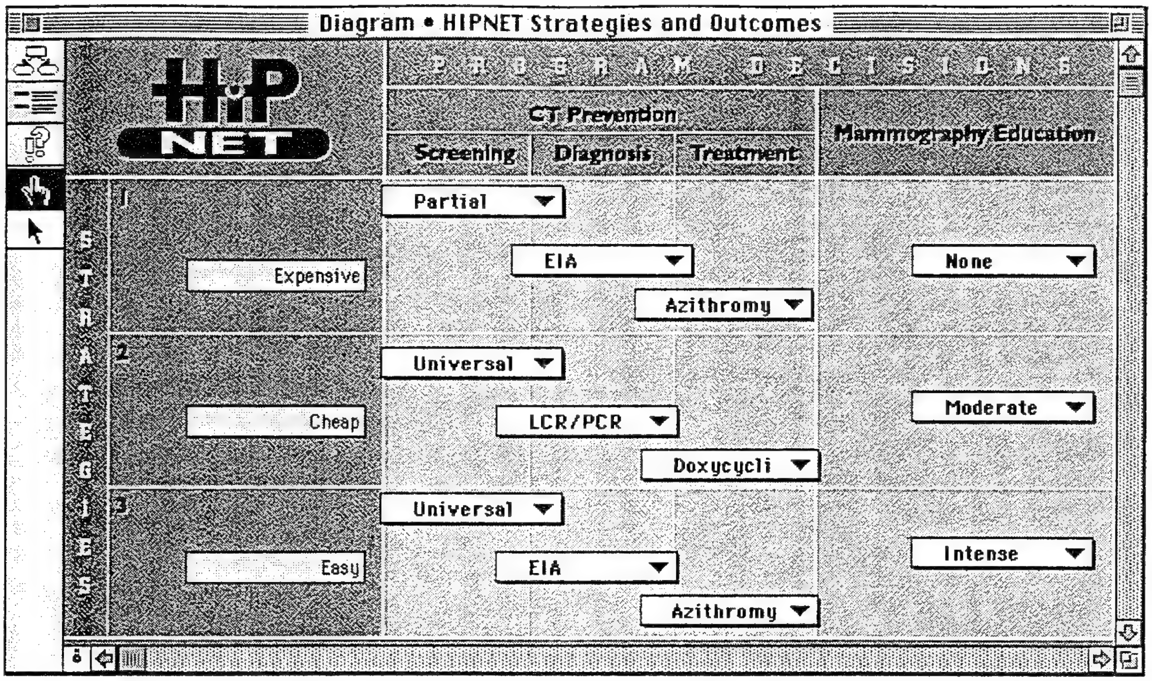


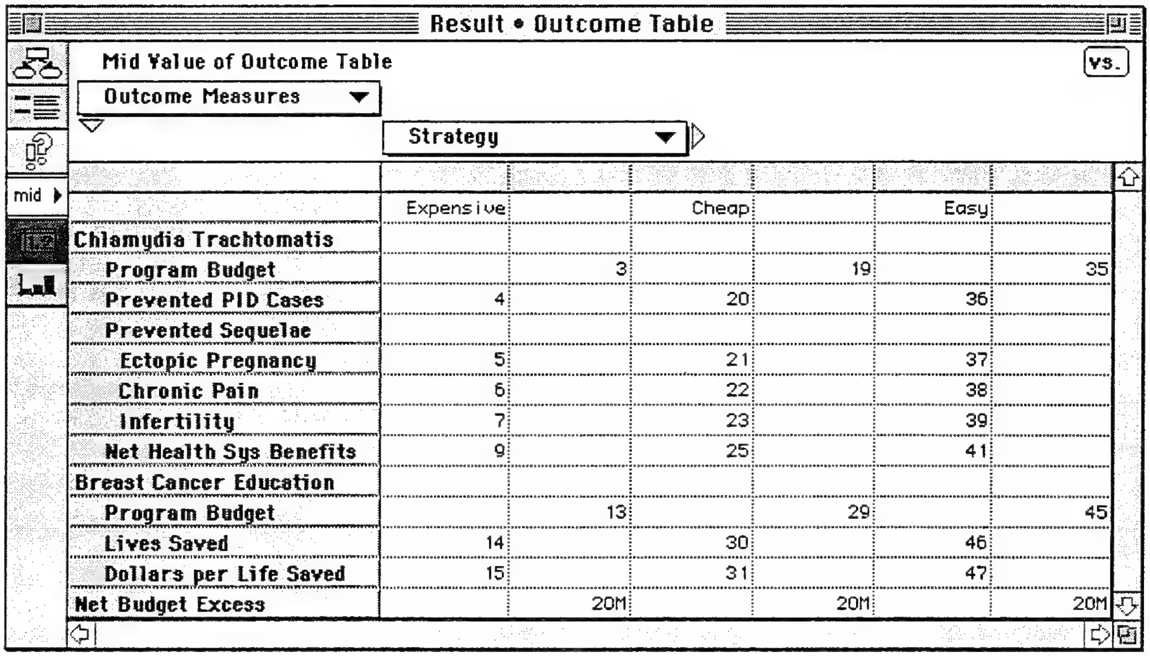
Figure 20: A strategy table to compare alternative strategies which fund programs at different levels. There is a popup menu for each program in each strategy that allows interactive selection of the policy option to define the characteristics of each of the three strategies. The Assembler combines the required IDMs to address the selected health programs.

Each HIPNet module is designed to help a public health decisions maker analyze a particular program or policy, such as Chlamydia prevention or Mammography education. Each option, for example choice of which population groups to screen, what diagnostic tests to perform, or what treatments to use will have a different distribution of costs and benefits. Decisions on one program, offered by a clinic or other healthcare agency, should not be examined in isolation, but in the context of other decisions, since many programs and policies may compete for the same scarce resources, especially funds and staff time.

A strategy table provides an integrated framework for decision makers to look at a set of decisions about a collection of programs or policies. Figure 20 shows an example strategy table for policies on Chlamydia prevention, including screening, diagnosis, and treatment, and mammography education. The strategy table allows the user to create and compare a number of different strategies — three in Figure 20. Each program and subprogram has a popup menu for each strategy with which the user can select the policy option for that strategy. For example, for the screening component of CT (Chlamydia Tracomatis) Prevention, the user can select partial or universal screening — that is, only high risk patients, or all clinic patients. For the diagnostic test, the user can select EIA or LCR/PCR. For the treatment antibiotic, the user can select Doxycyclin or Azithromycin. For the mammography education program, the user can select none, moderate, or intense policy options. The user can label the strategies — "Expensive", "Economical", and "Easy" in the example. The labels entered into

the table will automatically appear in tables and graphs comparing the strategies..

Once the user has defined a set of strategies by selecting options for each program for each strategy, HIPNet will compute a comparison table, as shown in Figure 21. The comparison table evaluates each strategy according to a variety of criteria, including effectiveness in preventing diseases and their sequelae, and in saving lives, as well as their cost, in comparison to the program budgets. This strategy table allows tradeoffs between programs within an integrated plan.



	Expensive		Cheap		Easy	
Chlamydia Trachtomatis						
Program Budget		3		19		35
Prevented PID Cases	4		20		36	
Prevented Sequelae						
Ectopic Pregnancy	5		21		37	
Chronic Pain	6		22		38	
Infertility	7		23		39	
Net Health Sys Benefits	9		25		41	
Breast Cancer Education						
Program Budget		13		29		45
Lives Saved	14		30		46	
Dollars per Life Saved	15		31		47	
Net Budget Excess		20M		20M		20M

Figure 21: Comparison table that shows budgets by program and evaluates the selected strategies according to selected criteria.

5. Displaying assumptions and identifying model insights

The ultimate goal of decision analysis and decision support is to help decision makers reach and implement more effective decisions. Responsible decision makers will seldom trust "black box" recommendations — that is recommendations without clarity about the underlying assumptions and reasoning. Hence, clear communication of the assumptions and reasoning leading to new insights are essential to effective decision support.

HIPNet, and Demos, the software used to implement the decision models, provide a variety of facilities to help document, and communicate both qualitative structures, formulas, and numerical input assumptions. Demos also provides sensitivity analysis tools to compute and display what is important and why. In this section, we provide a brief overview of those features not already illustrated earlier in this report.

The hierarchical influence diagrams themselves, as illustrated in previous sections of this report, provide a clear way to visualize the qualitative structure of decision models — both the dependencies among the variables, as displayed by the influence arrows, and the hierarchical structure shown by the module hierarchy and outline windows.

5.1 Documentation of public variables

For appropriate reuse of modules, it is essential that the assumptions, meaning, and scope of application of the module is clearly documented so that users are not tempted to apply it inappropriately. Within Demos, each variable contains structured documentation, providing the title, description, units of measurement of the variable, as well as its definition, dimensions, inputs and outputs. The dimensions, inputs and outputs (if any) are computed directly by Demos from the definitions to guarantee consistency. Figure 22 shows an example documentation card for such a variable.

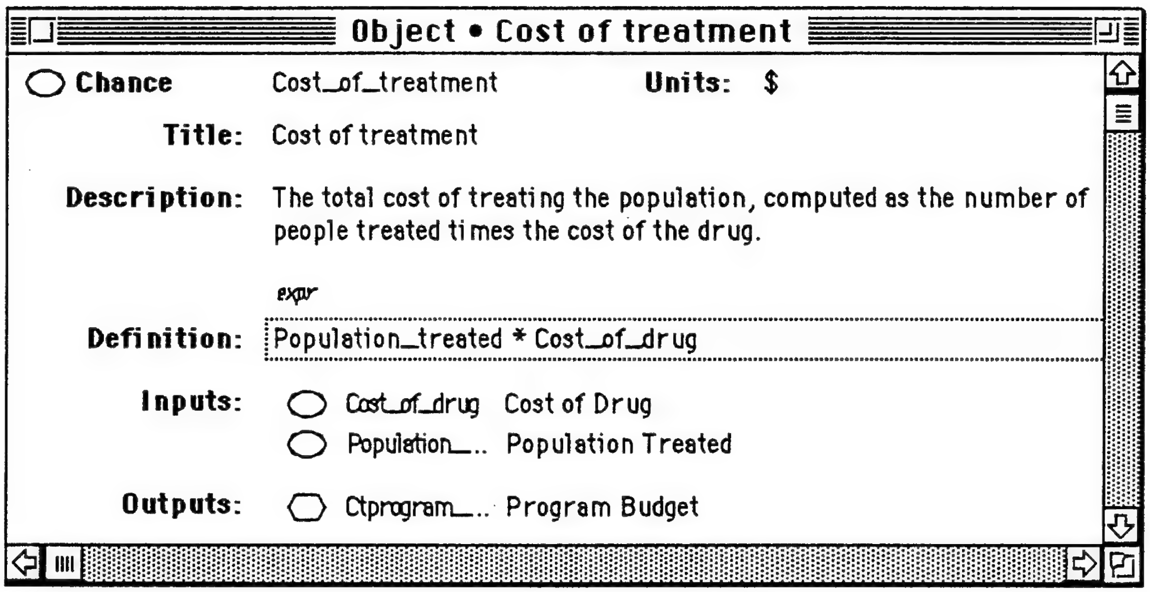


Figure 22: Documentation card attached to a variable, showing its units, description, definition, inputs, and outputs. The inputs are those variables appearing the definition, and are automatically kept consistent with the influences shown in the diagrams.

5.2 Model Recommendations

The recommendations of each decision model can be provided in a variety of formats. The strategy table of Figure 21 provides implicit recommendations in terms of the relative costs and benefits of each strategy option. HIPNet can provide individual graphs and tables for each decision, showing the costs, effectiveness, and benefits. Figure 23 shows one example with the cost-effectiveness of moderate and intense programs for education on the value of mammography, to persuade at-risk women to be screened for breast cancer.

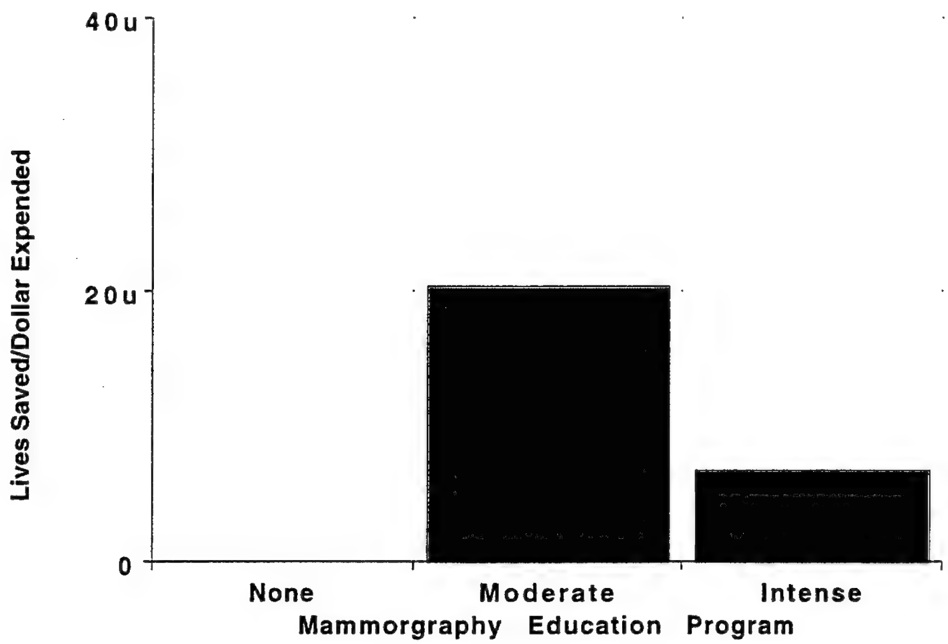


Figure 23. The **moderate** mammography program saves 20 micro-lives per dollar (20 lives per million dollars) expended versus about 6 micro-lives per dollar (6 lives per million dollars) for the **intense** program. The moderate program appears more cost-effective, although the intense program will save more lives.

5.3 Sensitivity and uncertainty analysis

Many numerical variables required for decision modules can be specified as uncertain, as described in Section 4.2. These uncertainties will give rise to uncertainty in the overall evaluations of the decision alternatives. Sensitivity analysis identifies the relative effect of each input variable on the results, for example the overall value of each option. Uncertainty analysis compares the relative contribution of sources of uncertainty in each input on the result. Figure 24 shows an example uncertainty analysis of the key uncertain inputs to the decision model for selecting an antibiotic treatment for Chlamydia. The "importance" is measured by the partial rank-order correlation of each input with respect to the output. The infection rate and test sensitivity are by far the most important variables according to this example. Therefore, if the decision maker wishes to improve the basis for the decision, it will be most cost-effective to obtain additional information about these quantities. For example, a more extensive survey to measure prevalence rate of Chlamydia, or a study of test sensitivity might be called for.

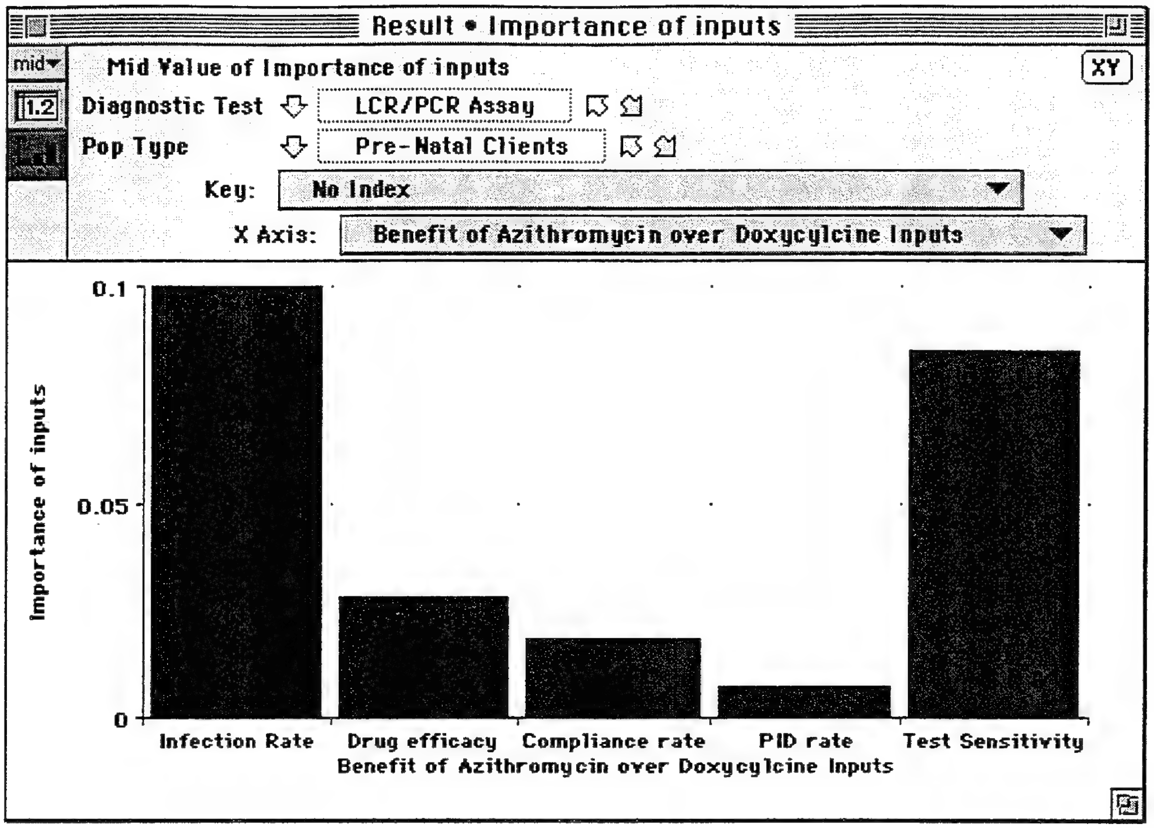


Figure 24: HIPNet Analysis Window for Chlamydia module showing the relative importance of key uncertain inputs. It shows that the infection rate of Chlamydia and the test sensitivity are the dominant contributors for this case.

6. Software architecture for DROL and HIPNet Prototype

The software architecture that we employed for the DROL and HIPNet prototype consists of three parts, using three different software environments:

1. The HIPNet introductory pages, IDM library directory, background information resources, and links to other public-health resources are implemented as **HTML** (HyperText Markup Language), the standard format for pages of the World Wide Web. They are therefore directly accessible via any Web browser.

2. The Adapter and Assembler programs that interview the user to obtain structuring and numerical information to adapt each module to the specifics of the user's application, are implemented in **Hypercard**, a widely used Macintosh tool for rapid prototyping.

3. The HIPNet modules, sensitivity analyses, strategy table, and comparison tables are implemented in **Demos**, the Decision Modeling System developed by Lumina, predecessor to Analytica. Using Demos, we were able to use hierarchical influence diagrams to create and test the decision models with speed and convenience. Demos provides rapid probabilistic analysis of the prototype models, and a wide variety of flexible displays for tables and

graphs of results. It also provides support for the integrated documentation discussed above.

We designed the user interfaces for all three kinds of components of HIPNet — in HTML, Hypercard, and Demos — using similar visual design, logo, and colors, so that they appear to the user as parts of an integrated environment. The choice of implementation environments was designed for convenience of rapid prototyping as part of the Phase I project, so that we could get early feedback on the essential concepts.

The full use of HIPNet requires execution of a Web browser, and downloading and execution of Hypercard and Demos components. This download takes a few minutes with 14,400 baud modem, including expanding of the compressed files. The Hypercard browser and Demos application can, if required, also be automatically downloaded via the Internet to the user's Macintosh, which takes a little longer. But, this would only need to be performed once for any HIPNet user. In this way, the user does not have to worry about setting up his or her own computer or obtaining software prior to using HIPNet.

7. Conclusions on achievements and future work

The goal of this Phase I SBIR research was to design innovative methods, and develop prototypes sufficient to illustrate and evaluate the feasibility and potential value of these methods. In the follow subsection, we summarize the achievements of the Phase I research, in response to each of the objectives from our Phase I proposal. In the final subsection, we outline directions in which this research might be developed in a future project designed to take these tools towards a generally usable, and eventually commercial, series of products.

7.1 Objectives and achievements

The following paragraphs restate *in italics* the original objectives of this work, followed by a brief summary of our achievements in meeting those objectives.

1. *The use of influence diagrams as an intuitive graphical method for creating and communicating extensible, reusable modules:* We have designed a number of extensions to the conventional influence diagram notation and semantics to support organizing complex models as a hierarchy of influence diagram modules and to facilitate definition, checking, and linking of these modules. We have successfully implemented and tested many of these extensions within Demos and Analytica, as part of our DROL prototype.

2. *A library facility for storage and retrieval of these modules, accessible world-wide via the World Wide Web:* We have designed and implemented a prototype Decision Resources On-line (DROL), with a small number of prototype IDMs, links to related information, and the ability to download selected IDMs.

3. *Tools to identify, adapt, and assemble these modules to create new models:* We have elaborated the concept and design of Adapters and Assemblers to help end users modify IDMs for application at a particular site for a particular situation, The Adapters can reconfigure the structure of a module, as well as requesting site-specific numerical values and distributions, in response to the users assessment of the local situation. The Assembler helps the user select and combine multiple IDMs to allow comparison of the costs and benefits of several strategies for adoption of public-health programs to be considered in the light of overall budget and other resource constraints. We have developed prototype Adapters for two IDMs and an Assembler to create a strategy table and evaluation combinations of policies.
4. *Communicate the assumptions, and key sensitivities to provide qualitative insights into model recommendations:* The prototype IDMs contain substantial structured explanatory material to describe and explain the variables, relationships, and assumptions in each model. The Adapter also provides a summary of key assumptions for each module. The analyzer provides sensitivity analysis to show parametrically how key parameters can change the recommended decision, and importance analysis to compare different sources of uncertainty. At the suggestion of Dr. Tom Garvey, our program director, at the inception of our project, we cut back on the original proposal to provide automated explanation facilities, in favor of more effort on developing practical applications.
5. *The application of these tools to an important class of decision problems:* We developed a prototype application of DROL to public health as HIPNet (Health Investment Planning Network), including decision modules for:
   - Public information programs for breast cancer screening
   - Chlamydia screening, testing, and treatment, with sample application to real problem to select the most cost-effective antibiotic treatment for Los Angeles County Department of Health Services
   - Measles vaccination program
   - Tuberculosis screening and treatment

In summary, we believe that we have achieved all the objectives of our original proposal, with the exception of the automated explanation an aspect of objective 4, which we did not attempt at the direction of ARPA.

We made considerably more progress in several areas than originally anticipated in our proposal. In particular, we developed a working prototype application for HIPNet, including many features that we had anticipated that we would only design in the Phase I of this research. Of particular note is the fact that Lumina has implemented several of the features designed to support modularity of influence diagrams developed in this project within Lumina's Analytica™ software, now in beta testing and scheduled for commercial release in summer 1996. While this step is far from constituting the commercialization of DROL as a

whole, it does demonstrate that several of the techniques developed under this research have immediate commercial applicability.

7.2 Directions for future work

Detailed delineation of our plans for future work based on this Phase I project will be provided in a proposal for Phase II research. Here we outline some promising directions:

- Identify of other promising classes of decision domains to demonstrate and evaluate the potential of libraries of reusable IDMs
- Improve support for multiattribute decision support for comparing attribute tradeoff weights, analyzing, and rating options for multiple collaborating decision makers at different sites,
- Develop IDM libraries based on a relational database, accessible via the Web, to provide rapid search, storage, security, access control, and version control that will scale up for large libraries
- Improve representations and tools for creating and editing IDMs, with features identified in this report, including private variables, access control, name scoping, units checking, index checking, graphical linking to public variables, and virtual libraries.
- Enhance features of object-oriented decision models, including improved inheritance, specialization of methods, extended variable types, database access variables, and remote Web access supporting OLE, CORBA or other cross-platform object access protocols.
- Development of translators between the IDMs used in Demos and Analytica and some other influence-diagram representations used by other software.
- Improve tools for creating Adapter and Assembler programs, supporting localization with site specific data obtained from networked databases where available.
- Develop DROL components, including library, Adapters, Assemblers, and model using more sophisticated user-interface and interaction with animation and direct manipulation made possible with add-ons for World Wide Web browsers, scripting, such as Livescript and JAVA. Use of these facilities will provide greater interactivity and speed without having to download sizable applications
- Provide automated textual and graphical explanation for decision models, decision recommendations, and sensitivities for ease of use by decision makers with no decision-analysis training.

Acknowledgments

We were assisted by numerous individuals and organizations in the project. We particularly wish to thank the following: Drs. Tom Garvey and Linda Harris of ARPA; Dr. Kevin Patrick, of the US Public Health Service, Washington, DC and University of California in San Diego; Drs. Steven Teutsch, Anne Haddix, Susan Hillis of CDC, Atlanta, GA; Drs. Janet Maylone, Arthur Thacher & Martin Wasserman of the Department of Health, Prince George's County, MD; Drs. Caswell Evans, Jonathan Freedman, and Gary Richwald, of Los Angeles County, Department of Health Services. Any errors in the article are the responsibilities of the authors.

The Lumina team that worked on this project included Max Henrion (Principal Investigator), Bob Korsan (Co-Principal Investigator), John Corbett, Brian Sterling, Rishi Varma, Brian Arnold.

References

Booch, G. (1991) *Object-oriented design: with Applications,* Benjamin Cummins, Redwood City, CA.

Breese, J. (1990) Construction of Belief and Decision Networks, Tech Memo 30, Rockwell International Science Center, Palo Alto, CA.

Fung, R. & Henrion, M. (1993). *Final report for integrated risk analysis for schedule and cost* (Final Report No. TR-93-01). Lumina Decision Systems.

Henrion, M, Breese, J.S. and Horvitz, E.J. "Decision Analysis and Expert Systems", Artificial Intelligence Magazine, Vol 12, No 4, Winter 1991, pp. 64-91.

Holtzman, S. (1988) *Intelligent Decision Systems,* Reading, MA: Addison-Wesley.

Howard, R. A. (1966). Decision analysis: applied decision theory. In D. B. a. M. Hertz J. (Ed.), *Proceedings of the Fourth International Conference on Operational Research,* (pp. 55-77).

Howard, R. A. (1968). The foundations of decision analysis. *IEEE-SMC,* 4, 211-219.

Howard, R. A., & Matheson, J. E. (1981). Influence diagrams. In R. A. Howard & J. E. Matheson (Eds.), *Readings on the Principles and Applications of Decision Analysis* (pp. 721-762). Menlo Park, Ca.: Strategic Decisions Group.

Howard, R. A., & Matheson, J. E. (Ed.). (1984). *Readings on the Principles and Applications of Decision Analysis.* Menlo Park, Ca.: Strategic Decisions Group.

Iman, R. L., & Helton, J.C. (1988). An Investigation of Uncertainty and Sensitivity Analysis Techniques for computer Models. *Risk Analysis,,* 8(1), 71-90.

Kahneman, D., Slovic, P., & Tversky, A. (Ed.). (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Keeney, R. L. (1977). The art of assessing multi-attribute utility functions. *OBHDP, 19,* 267-310.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value tradeoffs.* New York: Wiley and Sons.

Piscitello, D.M. & Chapin, A.L (1993) *Open Systems Networking,* Addison-Wesley, Reading, MA.

Raiffa, H. (1968). *Decision analysis: Introductory Lectures on Choice under Uncertainty.* Reading, Ma.: Addison-Wesley.

Shachter, R.D. (1986). Evaluating influence diagrams. In *Operations Research* (pp. 871-882).

von Winterfeldt, D., & Edwards, W. (1986). *Decision Analysis and Behavioral Research.* New York: Cambridge University Press.

Wellman, M.P. and Henrion, M., (1993) "Explaining 'Explaining Away'", IEEE *Transactions on Pattern Analysis and Machine Intelligence,* vol 15, No 3, March, 287:291.

Wiecha, C. & M. Henrion, (1988) "A graphical design environment for quantitative decision models", in *Expert Systems for Engineering Design*, M. Rychener (ed.), Academic Press, Boston, Mass, , pp 197-218.

Glossary

Adaptor	A program or agent that helps the user adapt a reusable influence diagram module for a specific situation
Analytica™	A software package developed by Lumina for decision modelling, the successor to Demos, due to be released in 3rd quarter 1996.
Assembler	A program or agent that helps the user assemble and link a set of influence diagram modules to address a specific decision problem
ARPA	Advanced Research Projects Agency of the US Department of Defense, the sponsor of this project.
CDC	The Centers for Disease Control and Prevention, an agency of the US Public Health Service, based in Atlanta, GA
CT	Chlamydia trachomata, a sexually transmitted disease
Demos™	Decision Modeling System, a commercially available software package developed and published by Lumina providing influence diagrams and probabilistic decision analysis, used as part of DROL.
DROL	Decision Resource Online Library
HIPNet	Health Investment Planning Network, a prototype decision resource for public health applications used to illustrate the tools developed in this project.
HTML	HyperText Markup Language, the standard format for pages on the World Wide Web
Influence diagram	A graphical representation of part of a decision model showing the variables and their qualitative dependences.
IDM	Influence-Diagram Module, a component of a decision model represented as an influence diagram, and structured as a module so that it can be reused for other applications.
Lumina	Lumina Decision Systems, Inc, the small business that performed the work described here.
SBIR	Small Business Innovation Research, the ARPA program under which this research was supported.
WWW	The World Wide Web